
 RELEASE

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSrch_n n a - n.a. database search, using Smith-Waterman algorithm
 Run on: Mon Mar 15 10:55:31 1999. MasPar time 7823.77 Seconds
 1545 001 Million cell updates/sec
 Tabular output not generated.

File: >US-09-020-716-5
 Description: (15115) from US09020716 seq
 Perfect Score: 5115
 N.A. Sequence: 1 GTTGGAGCTCTCCCATATG TCAAGCTATGATCCACAGC 5115
 Comp: CAACCTCGAGATGGATATAG AGTTCGATACCTAGGTTCCG

Scoring table: TABLE default

Gap 5

Mismatch STD: Dbase 0; Query 0

Searched: 602357 seqs, 118159623 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb157

1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
 7:em_om 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
 13:em_ro 14:em_un 15:em_vl

Database: genbank110
 16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
 22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl2 26:gb_pr
 27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
 33:gb_un 34:gb_vl

Statistics: Mean 12.348; Variance 5.557; scale 2.220

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2595	50.8	3000	32	CVGEM72FP Cloning vector pGEM-72	0.00e+00
2	2595	50.8	3003	32	CVGEM52FP Cloning vector pGEM-52	0.00e+00
3	2596	50.8	4486	32	EVOR116N Expression vector pCOR	0.00e+00
4	2598	50.2	3033	32	CVGEM7LICF Ligation-independent c	0.00e+00
5	2598	50.2	3033	32	CVGEM7LICF Ligation-independent c	0.00e+00
6	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
7	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
8	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
9	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
10	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
11	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
12	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00
13	2598	50.2	7823	32	CVGEM7LICF Ligation-independent c	0.00e+00

C 14	2200	43.0	4933	32	CVGEM52FP	Cloning vector pGEM-52	0.00e+00
C 15	2176	42.5	2743	32	CVGEM3Z	Cloning vector pGEM-3Z	0.00e+00
C 16	2176	42.5	4283	22	163120	Sequence 1 from patent	0.00e+00
C 17	2176	42.5	4283	22	185496	Sequence 1 from patent	0.00e+00
C 18	2153	42.1	6242	22	A25929	Yeast integration vect	0.00e+00
C 19	2148	42.0	6160	16	EC17DN3	E.coli pT7high-pl DNA	0.00e+00
C 20	2148	42.0	8953	16	EC17DN2	E.coli pT7high-pl DNA	0.00e+00
C 21	2148	42.0	11371	16	EC17DN1	E.coli pT7high-pl DNA	0.00e+00
C 22	2142	41.9	2686	32	XXU17492	Cloning vector pUC-Sce	0.00e+00
C 23	2142	41.9	2746	32	CVGEM42	Cloning vector pGEM-42	0.00e+00
C 24	2142	41.9	2364	32	SYNRLSKMV	BlueScript SK Minus cl	0.00e+00
C 25	2142	41.9	2964	32	SYNRLSKMV	BlueScript SK Minus cl	0.00e+00
C 26	2142	41.9	3811	32	SYNPS19V	PS19 cloning vector	0.00e+00
C 27	2144	41.9	3394	32	AF083409	Cloning vector p4S-lp	0.00e+00
C 28	2144	41.9	3394	32	AF083408	Cloning vector p4S-lp	0.00e+00
C 29	2142	41.9	3563	32	XXU04610	Cloning vector p4S-Tp	0.00e+00
C 30	2144	41.9	3604	32	AF083407	Cloning vector p4S-Cm	0.00e+00
C 31	2144	41.9	3604	32	AF083407	Cloning vector p4S-Cm	0.00e+00
C 32	2144	41.9	3617	32	AF083407	Cloning vector p4S-Cm	0.00e+00
C 33	2144	41.9	3714	32	AF083408	Cloning vector p4S-Gm	0.00e+00
C 34	2144	41.9	4236	32	AF083409	Cloning vector p4S-Km	0.00e+00
C 35	2144	41.9	4236	32	AF083409	Cloning vector p4S-Km	0.00e+00
C 36	2142	41.9	4312	32	AF083407	Cloning vector p4S-Tp	0.00e+00
C 37	2144	41.9	4503	32	AF083408	Cloning vector p4S-Sm	0.00e+00
C 38	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 39	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 40	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 41	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 42	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 43	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 44	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00
C 45	2142	41.9	5154	32	XXU17502	Cloning vector p4S-lp	0.00e+00

ALIGNMENTS

RESULT	1	CVGEM72FP	3000 bp	DNA	SYN	12-FEB-1996
LOCUS		Cloning vector pGEM-72f(+)				
DEFINITION		X65310				
ACCESSION		X65310				
NID		958180				
KEYWORDS		Beta-lactamase, bla gene, cloning vector, lacZ gene; Multiple cloning site; phage fl region; promoter.				
SOURCE		unidentified cloning vector				
ORGANISM		unidentified cloning vector				
REFERENCE		artificial sequence, cloning vectors.				
AUTHORS		1 (bases 1 to 3000)				
TITLE		Technical Services.				
JOURNAL		Direct Submission				
REMARK		Submitted (23-MAR-1992); Technical Services, Biomega Corporation.				
AUTHORS		2 (bases 1 to 3000)				
JOURNAL		Submitted (28-MAY-1993) Technical Services, Biomega Corporation.				
COMMENT		2800 Woods Hollow Road, Madison, WI 53711-5399, USA				
		This vector can be obtained from Biomega Corporation, Madison, WI.				
		Call one of the following numbers for order or technical information:				
		Order or technical 800-356-9526				
		In Wisconsin 800-356-9526				
		Outside U.S. 608-274-4330				
		See <X65310-X65335> for related vector sequences.				
FEATURES		Location/Qualifiers				
Source		1..3000				
		/organism="Cloning vector"				
		/plasmid="Plasmid pGEM"				
misc_feature		1				
		/note="7 transcription initiation site"				
misc_feature		10..110				
		/note="Multiple cloning sites"				
promoter		118..140				


```

Db 1718 CCCATGTTGTGCAAAAAGGGGTTAGTCTCTTGGTCTCCCGATCTGTCAGAAAGTAA 1777
      |||
Cp 3495 CCCCAGTGTGTCACAAAAGGAGGTTAGTCTCTTGGTCTCCCGATCTGTCAGAAAGTAA 3436
      |||
Db 1778 GTTGGCCGACAGTGTATACATCAAGGTTATGAGTACATCAAGTGTCTTACTGTCTAT 1837
      |||
Cp 3435 GTTGGCCGACAGTGTATACATCAAGGTTATGAGTACATCAAGTGTCTTACTGTCTAT 3376
      |||
Db 1838 GGCATCGGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 1897
      |||
Cp 3375 GGCATCGGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 3316
      |||
Db 1894 GGCATCGGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 1957
      |||
Cp 3315 GGCATCGGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 3256
      |||
Db 1958 TAGCAGAACTTTAAAGTGTCTATCTATGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 2017
      |||
Cp 3255 TAGCAGAACTTTAAAGTGTCTATCTATGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 3196
      |||
Db 2018 GATCTTACGCTGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 2077
      |||
Cp 3195 GATCTTACGCTGTGATGATGCTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 3136
      |||
Db 2078 AGCATCTTTTACTTTCACGAGGTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 2137
      |||
Cp 3135 AGCATCTTTTACTTTCACGAGGTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 3076
      |||
Db 2138 AAAAAGGGAATAGGGGACGACGGAATGTTGAATCTATCTCTCTCTCTCTCTCTCTCTCTCT 2197
      |||
Cp 3075 AAAAAGGGAATAGGGGACGACGGAATGTTGAATCTATCTCTCTCTCTCTCTCTCTCTCTCT 3016
      |||
Db 2198 TTATTGAAGCATTTATCAGGCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2257
      |||
Cp 3015 TTATTGAAGCATTTATCAGGCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2956
      |||
Db 2258 GAAATATAATTAATAGAGGTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 2317
      |||
Cp 2955 GAAATATAATTAATAGAGGTTTCTGTGAGTGTGAGTACATCAAGTGTCTTACTGTCTAT 2896
      |||
Db 2318 TGTGAATATACGACAGATGCTTGAAGGAGAAATACCGCATCAGGCGAAATTTGTAACGT 2377
      |||
Cp 2895 TGTGAATATACGACAGATGCTTGAAGGAGAAATACCGCATCAGGCGAAATTTGTAACGT 2836
      |||
Db 2378 TAAATTTTGTAAATTTCCGGTTTAAATTTTGTAAATCAGCTCATTTTTTAAACCAATA 2437
      |||
Cp 2835 TAAATTTTGTAAATTTCCGGTTTAAATTTTGTAAATCAGCTCATTTTTTAAACCAATA 2776
      |||
Db 2438 GGCAGAAATCGCAAAATTCCTTTATAATCAAAAGATAGACGAGATAGGTTGAGTGT 2497
      |||
Cp 2775 GGCAGAAATCGCAAAATTCCTTTATAATCAAAAGATAGACGAGATAGGTTGAGTGT 2716
      |||
Db 2498 TGTTCAGTTTGGCAAGAGTCCACTATTAAAGAGTGTGACTCCCACTCAAGGCGG 2557
      |||
Cp 2715 TGTTCAGTTTGGCAAGAGTCCACTATTAAAGAGTGTGACTCCCACTCAAGGCGG 2656
      |||
Db 2558 AAAACCGGTCTATCAGGCGGATGCCCCACTACGTGAACCATCAACCAATCAAGTTTTT 2617
      |||
Cp 2655 AAAACCGGTCTATCAGGCGGATGCCCCACTACGTGAACCATCAACCAATCAAGTTTTT 2596
      |||
Db 2618 GGGTGTGAGTGTGCTGTAAGGCTTAAATCGAAGCTTAAAGGAGCGCCGATTTAGAGC 2677
      |||
Cp 2595 GGGTGTGAGTGTGCTGTAAGGCTTAAATCGAAGCTTAAAGGAGCGCCGATTTAGAGC 2536
      |||
Db 2678 TTGAGGGGGAAGCGG 2693
      |||
Cp 2535 TTGAGGGGGAAGCGG 2520
      |||

```

```

RESULT 2
LOCUS CVR5M5ZFP 3003 bp DNA
DEFINITION Cloning vector pGEM-5zf(+).
          SYN 12-FEB-1996

```

```

X65308
q58176
bela-lactamase; bla gene; cloning vector; multiple cloning site;
phage fl region; promoter.
unidentified cloning vector.
unidentified cloning vector
artificial sequence; cloning vectors.
1 (bases 1 to 3003)
Technical Services.
Direct Submission
Submitted (23-MAR-1992) Technical Services, Promega Corporation.
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
revised by [2]
2 (bases 1 to 3003)
Klock, C.
Direct Submission
Submitted (28-MAY-1993) Technical Services, Promega Corporation.
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330
See <X65300-X65335> for related vector sequences.
FEATURES
source
1. 3003
/organism="Cloning vector"
/plasmid="Plasmid pGEM"
misc_feature
1
/notes="T7 transcription initiation site"
misc_feature
10..113
/notes="multiple cloning sites"
121..143
/notes="SP6 promoter"
126
/notes="SP6 transcription initiation site"
151..380
/notes="lac operon sequence"
complement(153..165)
/genes="lacZ"
/genes="lac2"
/notes="lac operator"
185..201
/notes="lac operator"
complement(1322..2192)
/genes="bla"
complement(1322..2182)
/genes="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/db_xref="PID:q58177"
/translation="MSIQHFRVALIPFFAFCPLVFAHPHETLVKVKDAEDLAGRVGY
IEDLNSGKILESFRPEEFPMVSTFKVLICHLIPALACEDQGPAPASQLIVE
YSPVTEKHLTDGNTVRELCSAAITMSDNTAANLLITIGGPKELTAFRNAGDHYRI
DRWEPELNEAIPNDERDITMPVAMPTITLTKLLTGELLTLASRQQLIDWWEACKVAGPT
LRSAIPAGFIADKSCAGGSPGIIAALGPFGKPSPIVVIYTTGSGQATMEBEPQIA
ETGASLIKHW"
misc_feature
2368..2823
/notes="phage fl region"
2824..2984
/notes="lac operon sequence"
2987..3003
/notes="T7 promoter"
770 a 743 c 764 g 726 t
BASE COUNT
ORIGIN
Query Match 50.8%; Score 2596; DB 32; Length 3003;
Best Local Similarity 100.0%; Pred No. of Gaps: 0;
Matches 2596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

101 GCTTGGATGATAGTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 160
115 GCTTGGATGATAGTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 5056
161 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 220
5055 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4996
221 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 280
4995 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4936
281 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 340
4935 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4876
341 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 400
4875 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4816
401 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 460
4815 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4756
461 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 520
4755 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4696
521 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 580
4695 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4636
581 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 640
4635 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4576
641 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 700
4575 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4516
701 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 760
4515 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4456
761 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 820
4455 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4396
821 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 880
4395 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4336
881 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 940
4335 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4276
941 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 1000
4275 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4216
1001 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 1060
4215 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4156
1061 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 1120
4155 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4096
1121 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 1180
4095 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 4036
1181 GTCATAGTGTGCTGTGAGTATGCTATAGTGTGACCTAAATAGCTGGGTAAATCATG 1240

4035 TGAAGCTCAGTGGAG 4976
1241 GATCTTCAGCTAGATGCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1400
3975 GATCTTCAGCTAGATGCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 4916
1301 TGAAGCTCAGTGGAG 1400
3915 TGAAGCTCAGTGGAG 4896
1361 CTGCTATTTCGCTCATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1420
3855 CTGCTATTTCGCTCATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4796
1421 GAGAGGCTTACGATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1480
3795 GAGAGGCTTACGATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4746
1481 TCCAGATTATACAGCAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1540
3735 TCCAGATTATACAGCAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4704
1541 AACTTTATCGGCTTCCATCCAGCTTATTAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1600
3675 AACTTTATCGGCTTCCATCCAGCTTATTAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 4616
1601 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1660
3615 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4596
1661 GTGCTTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1720
3555 GTGCTTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4496
1721 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1780
3495 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4446
1781 GTGCTTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1840
3435 GTGCTTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4404
1841 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1900
3375 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4344
1901 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1960
4315 GCGATTTAATAGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 4296
1961 TGAAGCTCAGTGGAG 2020
3255 TGAAGCTCAGTGGAG 4196
2021 GATCTTCAGCTAGATGCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 2080
3195 GATCTTCAGCTAGATGCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 4144
2081 AGCATCTTTTACTTTTACAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCT 2140
3135 AGCATCTTTTACTTTTACAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCT 4076
2141 AAAAAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAAT 2200
3075 AAAAAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAAT 4016
2201 TTATTGAGGCTTTTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2260
3015 TTATTGAGGCTTTTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2196
2261 AAAAAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAG 2320

Cp 2955 GAAAAATAAACAAATAGGGGTTCCGGGACACATTTCCCGGAAAAAGTGCACCTCTATCCGG 2896
 Db 2321 TGTGAATACCCGACAGATGGTATAGGAGAAATACCGCATCAGCGGAAATTTGTAACGT 2380
 Cp 2895 TGTGAATACCCGACAGATGGTATAGGAGAAATACCGCATCAGCGGAAATTTGTAACGT 2836
 Db 2381 TAATATTGTGTAATATCGGTTAAATATTGTTAAATCAGCTCATTTTTTAACCAATA 2440
 Cp 2835 TAATATTGTGTAATATCGGTTAAATATTGTTAAATCAGCTCATTTTTTAACCAATA 2776
 Db 2441 GCGGAAATCGCAAAATCCCTTAAATCAAAAGATAGACCCAGATAGGGTTGAGTGT 2500
 Cp 2775 GCGGAAATCGCAAAATCCCTTAAATCAAAAGATAGACCCAGATAGGGTTGAGTGT 2716
 Db 2501 TGTTCAGTTTGGAAACAGAGTCCACTATTAAAGAGGTGAGCTCCCAACGCTCAAGGGCG 2560
 Cp 2715 TGTTCAGTTTGGAAACAGAGTCCACTATTAAAGAGGTGAGCTCCCAACGCTCAAGGGCG 2656
 Db 2561 AAAAGCGGTCTATAGGCGGATGGGCGACTACGTGAACCATCAGCCCAAAATCAAGTTTTT 2620
 Cp 2655 AAAAGCGGTCTATAGGCGGATGGGCGACTACGTGAACCATCAGCCCAAAATCAAGTTTTT 2596
 Db 2621 GCGGTGAGGTGCGGTAAAGTCTAAATCGGAACCTTAAAGAGGTGAGCTCCCAACGCTCAAGGGCG 2680
 Cp 2595 GCGGTGAGGTGCGGTAAAGTCTAAATCGGAACCTTAAAGAGGTGAGCTCCCAACGCTCAAGGGCG 2536
 Db 2681 TTGACGGGGAAGCGG 2696
 Cp 2535 TTGACGGGGAAGCGG 2520

RESULT ?
 LOCUS EVCOR116N 4486 bp DNA SYN 24-MAR-1994
 DEFINITION Expression vector pCOR116N (modified from pCOR16 in [3]).
 ACCESSION Z29589
 NID q452349
 KEYWORDS ampicillin resistance; beta-lactamase; ColE1 origin; expression vector; nos terminator; phage fl region; rice actin1 promoter.
 SOURCE expression vectors.
 ORGANISM artificial sequence; expression vectors.
 REFERENCE 1 (bases 1 to 4486)
 AUTHORS Liu, L., Dasgupta, I., Davies, J., and Hull, R.
 TITLE Modified expression vectors for monocot transformation toward virus resistance
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4486)
 AUTHORS McElroy, D., Blowers, A. D., Jones, R. and Wu, P.
 TITLE Construction of expression vectors based on the rice actin 1 (Act1) 5' region for use in monocot transformation
 JOURNAL Mol. Gen. Genet. 231 (1), 150-160 (1991)
 MEDLINE 92092956
 REMARK (sites)
 REFERENCE 3 (bases 49 to 304)
 AUTHORS Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M.
 TITLE Napaline synthase: transcript and DNA sequence
 JOURNAL J. Gen. Appl. Microbiol. 1, 561-573 (1982)
 REFERENCE 4 (bases 1 to 4486)
 AUTHORS Liu, L.
 TITLE Direct Submission
 JOURNAL Submitted (26-JAN-1994) to the EMBL/GenBank/DBJ databases. Liu L., John Innes Institute, Virus Research, Colney Lane, Norwich, United Kingdom, NR4 7UH

FEATURES
 source
 Location/Qualifiers
 1..4486
 /organism="Expression vector"
 misc_feature
 1..52
 /note="multiple cloning sites"
 misc_feature
 3..5
 /note="start codon for gene expression"
 49..304
 /note="nos terminator"
 terminator
 /citation=[3]

misc_feature 511..967
 /note="phage fl region"
 CDS 1153..2013
 /note="beta-lactamase coding region"
 /codon_start=1
 /transl_table=11
 /db_xref="PID:q452349"
 /translation="MSIOHFRVALIPFFAAGCPVFAHPETLVKVKDAEDQLGARVSV
 IELDLNSGKILSFPPFPFMSTFKVLCHILLSPIDAGDGLGPARYSNDLVE
 YSVTEKHLDIGMTVELCSAALTMSDNTAANLLLTIGGKELTAFIRNKKDRVTR
 DRPELNEAIPNDERDITLKKLTIGELLGLASQQQLDKWEAKKVASPL
 LRSLALPGWFIADKSGAGSRGFIAGLPGDKPSPIVVIYITGSDAIMERNQA
 EIGASLIKHW"
 misc_feature 2014..2954
 /note="ColE1 origin"
 promoter 3256..4486
 /note="rice actin 1 5' region, (act1 promoter)"
 /citation=[2]
 BASE COUNT 1143 a 1118 c 1048 g 1177 t
 ORIGIN
 Query Match 50.8%; Score 2596; DB 32; Length 4486;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 2596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 639 CGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTT 698
 QY 2520 CGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTT 2579
 Db 699 ACGGACCTCGACCGCAAAAACCTTGATTTGGTGTATGGTTCAGTAGTGGGCTATCC 758
 QY 2580 ACGGACCTCGACCGCAAAAACCTTGATTTGGTGTATGGTTCAGTAGTGGGCTATCC 2639
 Db 759 CTGATAGACGGTTTTCGCTTTGAGCTTTGAGCTTCCAGTCTTAAATAGTGGAGCTTT 818
 QY 2640 CTGATAGACGGTTTTCGCTTTGAGCTTTGAGCTTCCAGTCTTAAATAGTGGAGCTTT 2699
 Db 819 GTTCCAAACTGGAAACACACTCAACCTATCTCGGCTATCTCTTTGATTTATAAGGAT 878
 QY 2700 GTTCCAAACTGGAAACACACTCAACCTATCTCGGCTATCTCTTTGATTTATAAGGAT 2759
 Db 879 TTTCGGGATTTTCGGCTTATGGTTAAAAATGAGCTGATTTAAACAAATTTAAACCGAA 938
 QY 2760 TTTCGGGATTTTCGGCTTATGGTTAAAAATGAGCTGATTTAAACAAATTTAAACCGAA 2819
 Db 939 TTTTAAACAAATATTAAGCTTTACAAATTCGCTGATGGGTATTTCTCTTACGCTC 998
 QY 2820 TTTTAAACAAATATTAAGCTTTACAAATTCGCTGATGGGTATTTCTCTTACGCTC 2879
 Db 999 TGTGGGTATTTACACCGCATACAGGTGACCTTTTCGGGGAATATGTGGGGGAACCC 1058
 QY 2880 TGTGGGTATTTACACCGCATACAGGTGACCTTTTCGGGGAATATGTGGGGGAACCC 2939
 Db 1059 TATTTGTTATTTTCTAAATATCAATATGATGATGATGATGATGATGATGATGATG 1118
 QY 2940 TATTTGTTATTTTCTAAATATCAATATGATGATGATGATGATGATGATGATGATG 2999
 Db 1119 ATAAATGCTTCAATATATGAAAAGGAAAGATGATGATGATGATGATGATGATGATG 1178
 QY 3000 ATAAATGCTTCAATATATGAAAAGGAAAGATGATGATGATGATGATGATGATGATG 3059
 Db 1179 CCTTATTCCTTTTTCGGGCAATTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1238
 QY 3060 CCTTATTCCTTTTTCGGGCAATTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3119
 Db 1239 GAAATTAAGATGCTGAAGATAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGG 1298
 QY 3120 GAAATTAAGATGCTGAAGATAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGG 3179
 Db 1299 CAACAGCGGTAAAGATCTCTGAGAGTCTTTCGCGCGGAAAGACGTTTCCCAATGATGAC 1358
 QY 3180 CAACAGCGGTAAAGATCTCTGAGAGTCTTTCGCGCGGAAAGACGTTTCCCAATGATGAC 2399


```

REFERENCE
AUTHORS      2 (bases 1 to 3033)
TITLE        Haun,R.S.
JOURNAL      Submitted (19-APR-1995) Randy S. Haun, NIH, NHLBI, Laboratory of
              Cellular Metabolism, Bldg. 10, Room 5K307, Bethesda, MD 20892-1434,
              USA
COMMENT
ligation-independent cloning vector. Preparation of the vector for
cloning includes linearization with Nari, gel purification of the
linearized vector, and treatment with T4 DNA polymerase in the
presence of dATP. The target sequence can be amplified using
sequence specific primers modified at the 5' end to contain an
additional 13 nucleotides complementary to the vector. The forward
primer should contain 5'-CIGCTCCGGCGA-3' followed by 12-15
nucleotide target-specific sequence. The reverse primer should
contain 5'-CIGCTCCGGCGA-3' followed by 12-15 nucleotide
target-specific sequence. The amplified sequence should be gel
purified and treated with T4 DNA polymerase in the presence of
dATP. Annealing of the vector and the amplification product forms a
duplex that can be used directly for transformation. Sequences
amplified using these primers are also compatible with the
p Bluescript II KS(+)/LIC vector (ATCC 87047) and pGEM-7zf(+)/LIC-F
vector (ATCC 87048). Differs from pGEM-7zf(+)/LIC-F (ATCC 87048)
only in the orientation of complementary ends generated at the
cloning site. This vector has been deposited in the ATCC repository
(ATCC 87049).
FEATURES
     source              Location/Qualifiers
     1..3033
     /organism="cloning vector pGEM-7zf(+)/LIC-R"
     /note="pGEM-7zf(+)/LIC-R. ATCC 87049"
     /db_xref="taxon:39794"
     misc_feature        55..81
     /evidence="experimental"
     /product="ligation-independent cloning region"
     65..70
     /evidence="experimental"
     /product="Nari restriction site"
     promoter            complement(151..173)
     /note="bacteriophage SP6 promoter"
     gene                complement(193..195)
     /gene="lacZ"
     CDS                 complement(193..195)
     /gene="lacZ"
     /note="initiation codon"
     /codon_start=1
     /transl_table=11
     /db_xref="PID:g806879"
     /translation="M"
     CDS                 complement(1952..2212)
     /codon_start=1
     /transl_table=11
     /product="bo3-lacZam3c"
     /db_xref="PID:g806880"
     /translation="MSIQHFRVALIPFFAAFLVFVAFHETLVKVKDAEDLGARVGY
     IELDLSGKILSFREPRFPMWSTFKVLGGLVSLRDAGQEQGLRIHYSONDLVE
     YSPVTEKHLDGWTVELCSAA.TMSDNTAANLLITTCGKELTAF.LHMGSDHVTRE
     DRPELPENALPNDERDITPVAMATTLKLLTGBELLASRQQLIDMWEADKVVAGEL
     LRSLAPAGNFIALKSONGERSGSGIIAALGDPKPSRVIVITTTGSOAIMDNERQIA
     EIGASLIKW"
     reprimin            complement(2398..2853)
     /standard_name="fl origin"
     promoter            join(3017..3033,1..6)
     /note="bacteriophage T7 promoter"
     BASE_COUNT          779 a 748 c 769 g 737 t
     ORIGIN
     Query Match          50.2% Score 2568; DB 32; Length 3033;
     Best Local Similarity 90.5%; Prod. No. 00e-00;
     Matches 2582; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
     Db 131 GGGTGGATCATAGTGTAGTATCTCTATAGTACCTAAATAGTGTGGCTAATCATG 190
     |||||||
     Cp 5115 GGGTGGATCATAGTGTAGTATCTCTATAGTACCTAAATAGTGTGGCTAATCATG 5056

```

```

Db 191 GCATAGCTGTTTCCTGCTGTAATTTGTTATGCGGCTCAAAATTCATCAAAATACAGAGC 250
|||||
Cp 5055 GCATAGCTGTTTCCTGCTGTAATTTGTTATGCGGCTCAAAATTCATCAAAATACAGAGC 4995
Db 251 CGAAGCATAAAGTGTAAAGGCTGGGGTCTCTAATGAGTGAAGTCAATCAGATTAATTCG 310
|||||
Cp 4995 CGAAGCATAAAGTGTAAAGGCTGGGGTCTCTAATGAGTGAAGTCAATCAGATTAATTCG 4936
Db 311 GTTGGCTCACTCCCGCTTTTCAGTTCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 370
|||||
Cp 4935 GTTGGCTCACTCCCGCTTTTCAGTTCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 4876
Db 371 CGGCAACGCTGCGGCGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 430
|||||
Cp 4875 CGGCAACGCTGCGGCGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4816
Db 431 TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
|||||
Cp 4815 TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4756
Db 491 AATACGGTTATCCACAGATCAAGGCGATCAAGGCGATCAAGGCGATCAAGGCGATCAAGGCG 550
|||||
Cp 4755 AATACGGTTATCCACAGATCAAGGCGATCAAGGCGATCAAGGCGATCAAGGCGATCAAGGCG 4695
Db 551 GCAAAAGGCTAGGAGGCTGTAAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
|||||
Cp 4695 GCAAAAGGCTAGGAGGCTGTAAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4636
Db 611 CCCTGACGAGCATCAACAAATTCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 670
|||||
Cp 4635 CCCTGACGAGCATCAACAAATTCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4576
Db 671 ATAAAGTACGAGCGCTTTCCGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
|||||
Cp 4575 ATAAAGTACGAGCGCTTTCCGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4516
Db 731 GCGCTTACGGGATACCTGTCGCGCTTTTCGCTTGGGAAAGCTGCTGCTGCTGCTGCTGCTGCT 790
|||||
Cp 4515 GCGCTTACGGGATACCTGTCGCGCTTTTCGCTTGGGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 4456
Db 791 CTCACGCTGAGGTATCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
|||||
Cp 4455 CTCACGCTGAGGTATCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4396
Db 851 CGAAGCCGCTTACGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
|||||
Cp 4395 CGAAGCCGCTTACGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4336
Db 911 CCGGCTTAAGACACACATTAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970
|||||
Cp 4335 CCGGCTTAAGACACACATTAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4276
Db 971 GAGTATGTAGGCGGTGTACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
|||||
Cp 4275 GAGTATGTAGGCGGTGTACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4216
Db 1031 AAGGACAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
|||||
Cp 4215 AAGGACAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4156
Db 1091 TAGCTCTGTAGTCGGGCAACAAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1150
|||||
Cp 4155 TAGCTCTGTAGTCGGGCAACAAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4096
Db 1151 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAATGCTTTCAGTCTTTCATGCGGCTGCTGCTG 1210
|||||
Cp 4095 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAATGCTTTCAGTCTTTCATGCGGCTGCTGCTG 4036
Db 1211 TGAGGCTTACGCGGCAACAAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1270
|||||
Cp 4035 TGAGGCTTACGCGGCAACAAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3976
Db 1271 GATCTTACCTAGATCTCTTAAATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1330

```



```

/misc_feature      /db_xref="taxon:39798"
55..81
/evidence-experimental
/product="ligation-independent cloning region"
56..71
/evidence-experimental
/product="Nari restriction site"
complement(151..173)
/Note="bacteriophage SP6 promoter"
complement(193..195)
/gene="lacZ"
CDS
complement(193..195)
/gene="lacZ"
/Note="initiation codon"
/codon_start=1
/transl_table=11
/db_xref="PID:g806890"
/translation="M"
complement(1352..2212)
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/db_xref="PID:g806891"
/translation="MSIQHREVALIPFAAFCLPVFAHPETLLVKKDAEDQLGARVGY
IEDLSKILLESERPERFPMSTFKVLICGLAVLSRIDAGQQLGRRIHYSONDLVE
YSPTEKHLIDGMVRELCSAAITMSDNTAANLLLTIGPKRELTAFLLHMGDHTPL
DRWEPELNEAIPNDRDTTPVAMAITLRKLLTGELTLASRQOLLDMWEADKVGPL
LRSLAPGWFTADKSGAGEFGSPGIIAALGPDGKPSRVIVVYITGSCATMDERNQIIA
EIGASLIKHW"
complement(2328..2853)
/standard_name="r1 origin"
join(3017..3033,1..5)
/Note="bacteriophage T7 promoter"
BASE COUNT      781 a 751 c 766 g 735 t
ORIGIN

Query Match      50.2%; Score 2568; DB 32; Length 3033;
Best Local Similarity 94.5%; Freq. No. 0.00+0.0;
Matches 2582; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

Dbb 131 GCCTTGATGATAGTTGAGTAATCTATAGTGCACCTAAATAGCTTGGGTAAATCATG 190
Cp 5115 GCGTTGATGATAGTTGAGTAATCTATAGTGCACCTAAATAGCTTGGGTAAATCATG 5056
Dbb 191 GTATATAGCTTTTCCGTGCTGAAATGTATATCCGCTCAAAATCCACACATACAGGC 250
Cp 5055 GTATATAGCTTTTCCGTGCTGAAATGTATATCCGCTCAAAATCCACACATACAGGC 4996
Dbb 251 CGGAATGATAAATGTAAAGGCTGAGGCTGAGTAAATGATGAGTAAATGATGAGTAA 310
Cp 4995 CGGAATGATAAATGTAAAGGCTGAGGCTGAGTAAATGATGAGTAAATGATGAGTAA 4936
Dbb 311 GTTGCTGATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATG 370
Cp 4935 GTTGCTGATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATG 4876
Dbb 371 CGGCTGATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATG 430
Cp 4875 CGGCTGATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATGAGTAAATGATG 4816
Dbb 431 TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
Cp 4915 TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4756
Dbb 491 AATACGTTATACAGATATCAAGGATATACAGGATATACAGGATATACAGGATATACAG 550
Cp 4755 AATACGTTATACAGATATCAAGGATATACAGGATATACAGGATATACAGGATATACAG 4896
Dbb 551 GCAAAAGGCGAGAAACCGTTAAAGAGGCGGCTTGTGCTGCGGTTTTTCCATAGGCTCCG 610
Cp 4695 GCAAAAGGCGAGAAACCGTTAAAGAGGCGGCTTGTGCTGCGGTTTTTCCATAGGCTCCG 4636
Dbb 611 CCGTGAAGGATATCAAAAAATGAAAGGATATCAAAAAATGAAAGGATATCAAAAAATGAA 670

```



```

REFERENCE 3 (bases 1 to 3199)
AUTHORS Klock C.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330
See <X65300-X65335> for related vector sequences.
FEATURES
source
Location/Qualifiers
1..3199
/organism="Cloning vector"
/plasmid="Plasmid pGEM"
misc_feature
1
/Note="T7 transcription initiation site"
5..61
/Note="multiple cloning sites"
64..86
/Note="SP6 promoter"
69
/Note="SP6 transcription initiation site"
94..323
/Note="lac operon sequence"
complement(106..108)
/Note="lacZ"
/Note="lacZ start codon"
complement(106..108)
/Note="lacZ"
/Note="bla"
complement(1265..2125)
/Note="bla"
complement(1265..2125)
/Note="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/db_xref="PID:g58173"
/transl_table="MSI0HFRVALIPFFAFLPVPFAHPETLVKVKDAEDQLGARVGV
IELDNSGKILSFEPERPMWSTFKVLLCHTLISRLDAGOEOLGRARYSONDLVE
YSPVTEKHLLTDGTVBELCSAITSNDTANLLITITGGKEULTAFHNGDVRTEL
DRPELNEAIPNDEKDTTPMVAFTLLKRLTGELTLASRQQLIDWMEADKRVAGPL
LRSLPAGNFIADKSGAGERGSPGIIAAIGPDGPKSRIVYITFGSQATMDERNPQIA
EIGASLIKHW"
misc_feature
2564..3019
/Note="phage f1 region"
3020..3190
/Note="lac operon sequence"
3183..3199
/Note="T7 promoter"
BASE COUNT 779 a 803 c 801 g 816 t
ORIGIN
Query Match 43.0% Score 2200; DB J2; Length 3199.
Seq. Id. Similarity 100.0% Pred. No. 0.00e+00;
Matches 2200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 57 AGCTTGAGTATCTATAGTGCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTT 116
Cp 5102 AGCTTGAGTATCTATAGTGCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTT 5043
Db 117 GTGTGTGAATTTGTTATCGGCTCACAATTCACACACATACAGCGCGGAACATAAAG 176
Cp 5042 CCGTGTGCAATTTGTATCGGCTCACAATTCACACACATACAGCGCGGAACATAAAG 4983
Db 177 TGTAAAGCTGGGTGCTATAGTGAAGTACATACATATATTCGGTTCGCTCACTG 236
Cp 4982 TGTAAAGCTGGGTGCTATAGTGAAGTACATACATATATTCGGTTCGCTCACTG 4923
Db 237 CCGGTGTTTCTAATGCGGAAACCTGTTGTCGACGTGCAATTAATGAATCGGCCAACCGCG 296
Cp 4922 CCGGTGTTTCTAATGCGGAAACCTGTTGTCGACGTGCAATTAATGAATCGGCCAACCGCG 4863

```

```

Db 297 GGAGAGGGGGTTTGGTATTGGGGCTCTTCGGCTTCCTGGCTCAGTCTCGGTGGC 356
Cp 4862 GGAGAGGGGGTTTGGTATTGGGGCTCTTCGGCTTCCTGGCTCAGTCTCGGTGGC 4203
Db 357 TCGTTCGTTTCGGCTTCGGGGGAGAGGATATAGTCTCAATCAAAAGGAGGATATATCC 416
Cp 4802 TCGTTCGTTTCGGCTTCGGGGGAGAGGATATAGTCTCAATCAAAAGGAGGATATATCC 4743
Db 417 ACAGATCAGGGGATTAACGACGAGAAACATGTGAGCAAAAGGCGCAGCAAAAGGCGAGG 476
Cp 4742 ACAGATCAGGGGATTAACGACGAGAAACATGTGAGCAAAAGGCGCAGCAAAAGGCGAGG 4663
Db 477 AACCGTAAAGAGGCGCGTTTGGTGGGCTTTTTCGATAGGCTCCGCCCTCGACGAGCAT 536
Cp 4682 AACCGTAAAGAGGCGCGTTTGGTGGGCTTTTTCGATAGGCTCCGCCCTCGACGAGCAT 4623
Db 537 CACAAAATGACGCTCAAGTCAGAGTGGCGAAACCGGACAGGACTATAAGATATACAG 596
Cp 4622 CACAAAATGACGCTCAAGTCAGAGTGGCGAAACCGGACAGGACTATAAGATATACAG 4563
Db 597 GCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTTCTGACGCTGCTGCTGCTGCTG 656
Cp 4562 GCGTTTCCCGCTGGAAGCTCCCTCGTGGGCTCTCTGTTTCTGACGCTGCTGCTGCTGCTG 4503
Db 657 TACGTCTCCCGCTTTCGCTTCGGGAAAGCGTGGGCTTCTTCATAGCTCAAGCTGTATAG 716
Cp 4502 TACGTCTCCCGCTTTCGCTTCGGGAAAGCGTGGGCTTCTTCATAGCTCAAGCTGTATAG 4443
Db 717 TATCTCAGTTCCGTTGAGTCTGCTCCAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 776
Cp 4442 TATCTCAGTTCCGTTGAGTCTGCTCCAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGT 4383
Db 777 CAGCGGACCGCTCGGCTTATCCGCTTACCTATGCTCTTCTGAGTCTCAAGCTGTATAG 836
Cp 4382 CAGCGGACCGCTCGGCTTATCCGCTTACCTATGCTCTTCTGAGTCTCAAGCTGTATAG 4323
Db 837 GACTTATCGGCTTTCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 896
Cp 4322 GACTTATCGGCTTTCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4263
Db 897 GTGTCTCAGAGTTCTTGAAGTGGTGGCTTAACTAGGCTTACACTAGAGGACACATATT 956
Cp 4262 GTGTCTCAGAGTTCTTGAAGTGGTGGCTTAACTAGGCTTACACTAGAGGACACATATT 4203
Db 957 GGTATCTGCGCTCTGCTGAGGCGAGTTACCTCGGAAAGAGATTTGCTAGCTCTTATGATCC 1016
Cp 4202 GGTATCTGCGCTCTGCTGAGGCGAGTTACCTCGGAAAGAGATTTGCTAGCTCTTATGATCC 4143
Db 1017 GGCAACAAACACCGCTGGTACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1076
Cp 4142 GGCAACAAACACCGCTGGTACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4083
Db 1077 AGAAAAAAGGATCTCAAGAGATCTTTCGATCTTTTTCAGGGGCTTTCAGGCTCAGTGG 1136
Cp 4082 AGAAAAAAGGATCTCAAGAGATCTTTCGATCTTTTTCAGGGGCTTTCAGGCTCAGTGG 4023
Db 1137 AACGAAACCTCAGTTAAGGATTTTGGTTCATGAGATTAACAAAGAGATCTTTCAGGCTCAG 1196
Cp 4022 AACGAAACCTCAGTTAAGGATTTTGGTTCATGAGATTAACAAAGAGATCTTTCAGGCTCAG 3963
Db 1197 ATCCTTTTAAATTAAGTATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1256
Cp 3962 ATCCTTTTAAATTAAGTATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3903
Db 1257 TCTCAGAGTTACCAATGCTTTAATCAGTGGGACCTTATCTCAGGCTATCTGCTATTTGCT 1316
Cp 3902 TCTCAGAGTTACCAATGCTTTAATCAGTGGGACCTTATCTCAGGCTATCTGCTATTTGCT 3843
Db 1317 TCATCCATAGTTGCTGACTCCCGCTGCTGTAGTAACTACCATACGAGGAGGCTTACCA 1376
Cp 3842 TCATCCATAGTTGCTGACTCCCGCTGCTGTAGTAACTACCATACGAGGAGGCTTACCA 3783

```



```

58766
beta lactamase, bla gene, cloning vector, lacZ gene; multiple
cloning site; phage f1 origin; promoter, pUC/M13 primer binding
site.
SOURCE          UNIDENTIFIED CLONING VECTOR.
ORIGIN          UNIDENTIFIED CLONING VECTOR.
REFERENCE       ARTIFICIAL SEQUENCE; CLONING VECTORS.
AUTHORS         1. (bases 1 to 3223)
TITLE           Direct Submission.
COMMENT         Cite as: J. MAP (1993). Technical Services, Promega Corporation,
                2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARKS         revised by [2]
FEATURES        1. (bases 1 to 3223)
                 Key/C.C.
DIRECT SUBMISSION
SUBMITTED (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
This vector can be obtained from Promega Corporation, Madison, WI
Call one of the following numbers for order or technical
information:
order or Technical 800-356-9526
In Wisconsin    608-356-9526
Outside U.S.     608-274-4330
See Accession# X65345 for related vector sequences.
FEATURES        Location/Qualifiers
SOURCE          1..3223
                 /organism="cloning vector"
                 /plasmid="Plasmid pCERM"
misc_feature    1
                 /note="T7 transcription initiation site"
misc_feature    9..85
                 /note="multiple cloning sites"
promoter        88..110
                 /note="SP6 promoter"
misc_feature    93
                 /note="sp6 transcription initiation site"
misc_feature    118..137
                 /note="lac operon sequence"
misc_feature    128..144
                 /note="pUC/M13 reverse primer binding site"
misc_feature    complement(130..132)
                 /gene="lacZ"
                 /note="lacZ start codon"
                 /complement(130..132)
                 /gene="lacZ"
misc_feature    152..168
                 /note="lac operator"
gene            complement(1289..13043)
                 /gene="bla"
misc_feature    complement(1289..13043)
                 /gene="bla"
                 /note="interrupted reading frame"
                 /product="Beta-lactamase"
                 /complement(2149..2588)
                 /gene="bla"
misc_feature    complement(3044..3204)
                 /note="phage t1 region"
                 /complement(3044..3204)
                 /note="lac operon sequence"
                 /complement(3164..3182)
                 /note="pUC/M13 forward primer binding site"
                 /complement(3207..3223)
promoter        3207..3223
                 /note="T7 promoter"
BASE COUNT      783 a   811 c   804 g   820 t
ORIGIN
Quality Match      43.0%; Score 2200; DB ID: Length 3223;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
db      81 ACTTCAGATTCTATATGCTACCTAAATAGCTGTGGCGTAATCAIGGTCATAGCTGT 140
100|||||
p      102 AACTTCAGATTCTATATGCTACCTAAATAGCTGTGGCGTAATCAIGGTCATAGCTGTT 5043

```

Db 1221 ATCTTTTAAATTAATAAATGAAGTTTAAATCAATTAAGTATATATAGTAAACTTGG 1280
 Cp 3962 ATCTTTTAAATTAATAAATGAAGTTTAAATCAATTAAGTATATATAGTAAACTTGG 3903
 Db 1281 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTTCGT 1340
 Cp 3962 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTTCGT 3843
 Db 1341 TCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAACTACGATACGGGAGGCTTACCA 1400
 Cp 3842 TCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAACTACGATACGGGAGGCTTACCA 3783
 Db 1401 TCTGCCCCCAATGCTGCAATGATAACCGAGAGACCCAGCTCAGCGCTCCAGATTATCA 1460
 Cp 3782 TCTGCCCCCAATGCTGCAATGATAACCGAGAGACCCAGCTCAGCGCTCCAGATTATCA 3723
 Db 1461 GCAATAAACAGCAGCGGGAAGGGCGGAGCGAGAGTGGTCTGCAACTTTATCCGCC 1520
 Cp 3722 GCAATAAACAGCAGCGGGAAGGGCGGAGCGAGAGTGGTCTGCAACTTTATCCGCC 3663
 Db 1521 TCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAGTTCGGCAGTTAATAGT 1580
 Cp 3662 TCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAGTTCGGCAGTTAATAGT 3603
 Db 1581 TTGGCAACGTTGTTGGCATTTGCTACAGGCATCGTGGTGTACGCTGCTGCTGATG 1640
 Cp 3602 TTGGCAACGTTGTTGGCATTTGCTACAGGCATCGTGGTGTACGCTGCTGCTGATG 3543
 Db 1641 GCCTTCATTCAGCTCCGCTCCCAAGCATCAAGCGAGTTCACATGATCCCGCATTTGTC 1700
 Cp 3542 GCCTTCATTCAGCTCCGCTCCCAAGCATCAAGCGAGTTCACATGATCCCGCATTTGTC 3483
 Db 1701 AAAAAAGCGTTAGTCTCTCGGTCCTCCGATCGTTGTGAGAGTAAAGTTGCGCCAGTG 1760
 Cp 3482 AAAAAAGCGTTAGTCTCTCGGTCCTCCGATCGTTGTGAGAGTAAAGTTGCGCCAGTG 3423
 Db 1761 TTATCAGTATGTTATGGCAGCAGCTGCATAATCTCTTACTGTGATGATCCGATTAAGA 1820
 Cp 3422 TTATCAGTATGTTATGGCAGCAGCTGCATAATCTCTTACTGTGATGATCCGATTAAGA 3363
 Db 1821 TGTCTTCTGCTGACTGTGAGTACTCAACCAAGTCAATCTGAGATACCGCGCCGGGA 1880
 Cp 3362 TGTCTTCTGCTGACTGTGAGTACTCAACCAAGTCAATCTGAGATACCGCGCCGGGA 3303
 Db 1881 CCGAGTTGCTCTTCCCGGGGCTCAATACGGGATAATAGTGTATGACATAGCAGAACTTTA 1940
 Cp 3302 CCGAGTTGCTCTTCCCGGGGCTCAATACGGGATAATAGTGTATGACATAGCAGAACTTTA 3243
 Db 1941 AAGTGTCTCATTTGGAACAGCTTCTTCCGGGCGGAAACTCTCAAGGATCTTACCGGTG 2000
 Cp 3242 AAGTGTCTCATTTGGAACAGCTTCTTCCGGGCGGAAACTCTCAAGGATCTTACCGGTG 3183
 Db 2001 TTGAGTCCAGTTGATGTAAACCACTCTGTCACCACTGATCTTCAGCATCTTTTACT 2060
 Cp 3182 TTGAGTCCAGTTGATGTAAACCACTCTGTCACCACTGATCTTCAGCATCTTTTACT 3123
 Db 2061 TTCACAGCGTTTCTGGGTGAGCAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATA 2120
 Cp 3122 TTCACAGCGTTTCTGGGTGAGCAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATA 3063
 Db 2121 AGCGGCAACGGAATTTGATTAATCTCATCTCTTCTCTTTTCAATATTATTGAAGCAT 2180
 Cp 3062 AGCGGCAACGGAATTTGATTAATCTCATCTCTTCTCTTTTCAATATTATTGAAGCAT 3003
 Db 2181 TATCAGGCTTATTGCTCATGAGCGGATACATATTGAAATGATTATAGAAAAATAACAA 2240
 Cp 3002 TATCAGGCTTATTGCTCATGAGCGGATACATATTGAAATGATTATAGAAAAATAACAA 2943
 Db 2241 ATAGGCTTATTGCTCATGAGCGGATACATATTGAAATGATTATAGAAAAATAACAA 3280
 Cp 2942 ATAGGCTTATTGCTCATGAGCGGATACATATTGAAATGATTATAGAAAAATAACAA 3203

RESULT 13
 LOCUS CVGEM112P 3223 bp DNA SYN 12-FEB-1996
 DEFINITION Cloning vector pGEM-112f(+).
 ACCESSION X65313
 NID 958168
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phage f1 region; promoter; pUC/M13 primer binding site.
 SOURCE unidentified cloning vector.
 ORGANISM unidentified cloning vector
 REFERENCE 1 (bases 1 to 3223)
 AUTHORS Technical Services.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation.
 REMARK 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 REVISION 2 (bases 1 to 3223)
 AUTHORS Klock, C.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation.
 COMMENT 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330
 See <X65300-X65335> for related vector sequences.
 FEATURES
 source
 1..3223
 /organism="Cloning vector"
 /plasmid="plasmid pGEM"
 misc_feature
 1
 /note="T7 transcription initiation site"
 misc_feature
 19..85
 /note="multiple cloning sites"
 promoter
 88..110
 /note="SP6 promoter"
 misc_feature
 93
 /note="SP6 transcription initiation site"
 misc_feature
 118..347
 /note="lac operon sequence"
 misc_feature
 128..144
 /note="pUC/M13 reverse primer binding site"
 gene
 complement(130..132)
 /gene="lacZ"
 misc_feature
 complement(130..132)
 /gene="lacZ"
 misc_feature
 152..168
 /note="lac start codon"
 gene
 complement(1289..2149)
 /gene="bla"
 CDS
 complement(1289..2149)
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /db_xref="PID:958169"
 /translation="MSIQHFRVALIPFFAFCPLVFAHPETLVKVKDAEDQLGARVGY
 IELDLSGKILLESFPEPPFPMSTFKVLLCHTLISRIDAGQQLGPAPYSQNDLVE
 YSPVTEKHLTDGMTAVRELCSAAITMSDNTAANLLLTIGGPKELTFLHNNKDHVRL
 DRWEPELNEATIPNDERDTTPVAMPPTLRKLLTGELTLASROOLDIMMEAKVAPL
 LPSALPASHWIALAKSSGAEFASPSIIIAALGPGGFPSPVIVITITGSATMEKRRQTA
 EIGASLIRHW"
 misc_feature
 2588..3043
 /note="phage f1 region"
 misc_feature
 3044..3204
 /note="lac operon sequence"
 misc_feature
 3164..3180
 /note="pUC/M13 forward primer binding site"
 promoter
 3207..3223


```

Cp 3062 AGGCGACACGGGAATGTTGAATACATACATACATCTCCCTTTTCAATATATTATGAGCATT 3003
Db 2181 TATCAGGTTATTGTCATGAGCGATACATATTTGATGATGATTTAGAAAATAAACAA 2240
|||||
Cp 3062 TATCAGGTTATTGTCATGAGCGATACATATTTGATGATGATTTAGAAAATAAACAA 2943
|||||
Db 2241 ATAGGGGTTCTGGGAAATATTTCTGAAAAGTGGCACTG 2280
|||||
Cp 2942 ATAGGGGTTCTGGGCAATTTCTGGGAAAAGTGGCACTG 2903
|||||

RESULT 14
LOCUS CVP6MLDC 4933 bp DNA SYN 13-FEB-1996
DEFINITION Cloning vector pSEM-luc.
ACCESSION X55316
NID 958213
KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; luc gene;
multiple cloning site; phage fl region; promoter; pUC/M13 primer
binding site.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4933)
AUTHORS Technical Services.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARKS revised by [2]
AUTHORS Klock, C.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330
See <X5300-X5533> for related vector sequences
FEATURES
Source 1 4933
/organism="Cloning vector"
/plasmid="Plasmid pSEM"
1
misc_feature 1
/notes="T7 transcription initiation site"
9..47
/notes="multiple cloning sites"
complement(57..1756)
/notes="luciferase cDNA"
complement(102..1754)
/genes="luc"
complement(102..1754)
/genes="luc"
/codon_start=1
/product="luciferase"
/db_xref="PID:958213"
/translation="MEDAKNIKKGPAPYPLEDGTAGBOLHKMKRYALVPGTIAFTD
ARFENITVAEYFEMSVRLAFAMKPYGLNTNHRIVVCSNSLOFEMPVLGALFQVAV
APANDIYNRELLNSMNSIQPTVFEVSKGLQKILNVOKKPLIQKLIWDSTIDYG
FSMYIFVLSHLPFGNFEDFVPFSDRDKTIALIMNSGSGTGLPKGVALPHRTACVP
FSHARDPIFGNOIIPDTAILSVVPHFGFMFTTLGYLIGCFRVLMYRFEELFRLS
LDQYKTSALLVPTLFSFAKSLIDKYLNLHEIASGGAPLSKEVGEAVAKRFLHP
GIRQGYLTETTSAILITPEGGDKPGKGVVPEFAKVVDLDTGKTLGVNORGLCV
RPMIMSGVNVNPEATNALIDKGMHSDIAYWDEHFFIVDRLSLIKLYKGYOVA
PAELSLIHPNIDAGVAGLPDDAGELPAVVVLEHGTWTEKEILVDYVASQVTT
AKLPGGVVFPDEVKPLGTGKLDARKIPEILIKAKKGSKL"
1762..1795
/notes="multiple cloning sites"
1778..1794
/notes="pUC/M13 reverse primer binding site"
1798..1820
misc_feature 1
misc_feature 1
promoter

```

```

/misc_feature 1803
/notes="SP6 promoter"
1803
/misc_feature 1826..2057
/notes="lac operon sequence"
complement(1840..1842)
/genes="lacZ"
/notes="lacZ start codon"
complement(1840..1842)
/genes="lacZ"
1853..1880
/misc_feature 1853..1880
/notes="lac operator"
complement(2999..3859)
/genes="bla"
complement(2999..3859)
/genes="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/db_xref="PID:958213"
/translation="MSIQHFRVLVPPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
TELDNSKDKILESPRPEPPPMSTKRVLLCHILLSPIDAGQGLAPRAVRSQNDLVE
YSPVTEKHLIDGNIVRELCSNAITMSDNTAANLLIIIGPKELFTALRNHMGHVEL
DRWEPELNEAIPNDERDITMPVAMPITLLRKLIIIGELLILASRQOLIDWMEADKAGPE
LRSLAPAGWFTADKSGAGRPSPGIIAALGPMPKPSPIVIVITIGSDATMPFPRQIA
EIGASLIKHW"
4298..4751
/misc_feature 4298..4751
/notes="phage fl region"
4754..4912
/misc_feature 4754..4912
/notes="lac operon sequence"
4874..4890
/misc_feature 4874..4890
/notes="pUC/M13 forward primer binding site"
4917..4933
promoter 4917..4933
/notes="T7 promoter"

BASE COUNT 1234 a 1217 c 1165 g 1317 t
ORIGIN
Query Match 43.0%; Score 2200; DB 32; Length 4933;
Best Local Similarity 100.0%; Pred. No. 0 00e+00;
Matches 2200; Conservative 0; Mismatches 0; Labels 0; Gaps 0;
Db 1791 AGCTTCAGTATTCTATAGTGCACCTAAATAGCTTGGCGGTAACTATGCTGCTGCTGCTT 1850
|||||
Cp 5102 AGCTTCAGTATTCTATAGTGCACCTAAATAGCTTGGCGGTAACTATGCTGCTGCTGCTT 5043
|||||
Db 1851 CTTGTGTGAAATTTTATCGCTCACATTCACACATACAGAGCGGGAAGCAIAAG 1910
|||||
Cp 5042 CTTGTGTGAAATTTTATCGCTCACATTCACACATACAGAGCGGGAAGCAIAAG 4983
|||||
Db 1911 TGTAAAGCGCTGGGTCGCTTAATAGTGAAGCTAACTCACCATTAAITGGCTTGGCTGCTG 1970
|||||
Cp 4982 TGTAAAGCGCTGGGTCGCTTAATAGTGAAGCTAACTCACCATTAAITGGCTTGGCTGCTG 4923
|||||
Db 1971 CCGGCTTTCCAGTCGGGAACCGTTCGTCGCCAGCTGCATTATGAAACGGCAACGGCGCG 2030
|||||
Cp 4922 CCGGCTTTCCAGTCGGGAACCGTTCGTCGCCAGCTGCATTATGAAACGGCAACGGCGCG 4863
|||||
Db 2031 GCGAGAGGCGGCTTTCGCTATTCGCGGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2090
|||||
Cp 4862 GCGAGAGGCGGCTTTCGCTATTCGCGGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 4803
|||||
Db 2091 TCGGCTGCTGCGGCTGGCGGAGCGGTATCAGGTACTCAAAAGGAGGAGGAGGAGGAGGAGG 2150
|||||
Cp 4802 TCGGCTGCTGCGGCTGGCGGAGCGGTATCAGGTACTCAAAAGGAGGAGGAGGAGGAGGAGG 4743
|||||
Db 2151 ACAGATCAGGCGGATACACGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2210
|||||
Cp 4742 ACAGATCAGGCGGATACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4683
|||||
Db 2211 AACCGTAAAAAGCGCGCTTGGTGGCGTCTTTTCGATAGCGCTGCGCGCGCGCGCTGACGAGCAT 2270
|||||
Cp 4682 AACCGTAAAAAGCGCGCTTGGTGGCGTCTTTTCGATAGCGCTGCGCGCGCGCTGACGAGCAT 4623
|||||

```

W P I R E H

(TW)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSEARCH n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 17:10:57 1999: MasPar time 725.31 Seconds
Tabular output not generated 959.461 Million cell updates/sec
Title: >US-09-020-716-5
Description: (1-5115) from US09020716.seq
Perfect Score: 5115
N.A. Sequence: 1 GTTGGGAGCTCTCCCATATG TCAAGCTATGATCAACGC 5115
Comp: CAACCTCGAGAGGTATAC AGTTCGATACGTAGTTGG

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 188442 seqs, 58026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 10.354; Variance 5.797; scale 1.786
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2195	42.9	10306	12	Q73734	0.00e+00
2	2195	42.9	10470	12	Q73735	0.00e+00
3	2172	42.5	4118	35	T69189	0.00e+00
4	2172	42.5	4118	35	T69188	0.00e+00
5	2176	42.5	4283	35	T86449	0.00e+00
6	2176	42.5	4283	35	T86450	0.00e+00
7	2142	41.9	3690	30	V14310	0.00e+00
8	2142	41.9	4539	14	Q87347	0.00e+00
9	2142	41.9	6295	39	V02043	0.00e+00
10	2141	41.9	6824	6	Q39050	0.00e+00
11	2141	41.9	7287	39	V02042	0.00e+00
12	2142	41.9	10930	20	Q91256	0.00e+00
13	2142	41.9	10950	20	Q91255	0.00e+00

14	2138	41.8	4145	7	Q40379	0.00e+00
15	2138	41.8	4540	32	T69891	0.00e+00
16	2139	41.8	4824	20	T08975	0.00e+00
17	2138	41.8	4883	4	Q25709	0.00e+00
18	2138	41.8	5178	27	T49876	0.00e+00
19	2138	41.8	5356	25	T43794	0.00e+00
20	2138	41.8	5534	25	T43137	0.00e+00
21	2139	41.8	5585	14	O87418	0.00e+00
22	2140	41.8	5618	15	O88310	0.00e+00
23	2138	41.8	5620	4	Q27489	0.00e+00
24	2140	41.8	5653	17	T03998	0.00e+00
25	2140	41.8	6207	7	Q43813	0.00e+00
26	2138	41.8	6216	37	V01425	0.00e+00
27	2138	41.8	6229	20	T29220	0.00e+00
28	2140	41.8	6231	30	T60737	0.00e+00
29	2138	41.8	6238	39	V14339	0.00e+00
30	2140	41.8	6274	7	Q43812	0.00e+00
31	2138	41.8	6338	37	T91384	0.00e+00
32	2139	41.8	6367	18	T12661	0.00e+00
33	2139	41.8	6367	27	T48803	0.00e+00
34	2140	41.8	6555	30	T70315	0.00e+00
35	2139	41.8	6926	18	T12662	0.00e+00
36	2138	41.8	7050	7	Q40419	0.00e+00
37	2140	41.8	7106	15	Q47494	0.00e+00
38	2140	41.8	7106	8	Q47193	0.00e+00
39	2140	41.8	7298	8	Q47192	0.00e+00
40	2138	41.8	7474	31	T71320	0.00e+00
41	2139	41.8	8710	17	Q86650	0.00e+00
42	2139	41.8	9020	17	T07652	0.00e+00
43	2138	41.8	10288	31	T71322	0.00e+00
44	2138	41.8	10443	31	T71319	0.00e+00
45	2138	41.8	16656	15	Q86154	0.00e+00

ALIGNMENTS

RESULT 1
ID Q73734 standard: DNA: 10306 BP.

AC Q73734;
DE Retro virus vector p521;
KW GalV, gibbon ape-leukemia virus, retro virus, vector, plasmid p558;
KW plasmid p521; plasmid p537; transgenic animal; gene therapy; ss.
NN Synthetic.
PD W09423048-A.
FN 13-OCT-1994.
PF 06-APR-1994; 003784.
PR 06-APR-1993; US-043311.
PA (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Deacon NJ, Eiden MV, Hooker DJ, Wilson CA;
DR WPI: 94-333206/41.
PT Recombinant DNA contg. defective gibbon ape leukaemia virus
PT genome - and specific gene, also derived mammalian cells, viruses
PT etc., useful in gene therapy of e.g. cancer, viral infections or
PT inherited diseases
PS Disclosure: Page 41-46; 68pp; English.
CC Novel replication-deficient retro virus vectors p558 (Q73737), p521
CC (Q73734) and p537 (Q73735) include a GalV-derived packaging site,
CC esp. the 200-910 base region of the GalV S8A0 genome (Q73731), and
CC regulatory sequences, partic. the GalV 3' long terminal repeat
CC (Q73732). The vectors allow transfer of genes for gene therapy and
CC transgenic animal breeding.
SQ Sequence 10306 BP: 2400 A; 2758 C; 2733 G; 2415 T;
Query Match 42.9%; Score 2195; E5 12; Length 10306;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2195, Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 8112 caggtggcacttttcggggaatgtatcgaggaacccatttttttttataaac 8171
QY 2903 CAGGTGGCAGCTTTTCGGGGAATGTCGAGGAACCCATTATTTTCTAAATAC 2962

ID Q73735 standard: DNA: 10970 BP.
 AC Q73735;
 AT 10-MAY-1995 (first entry)
 DE Retro virus vector p537
 KW GalV: gibbon-ape-leukemia virus; retro virus; vector; plasmid p558;
 KW plasmid p521; plasmid p537; transgenic animal; gene therapy; ss
 US Synthetic.
 PN WO9423048-A.
 PD 13-OCT-1994.
 PR 06-APR-1994; U03784.
 PR 06-APR-1993; US-043311.
 PA (USSH) US SEC DEPT HEALTH.
 PA (USSH) US NPPT HEALTH & HUMAN SERVICES.
 PI Deacon NJ, Eiden MV, Hooker DJ, Wilson CA.
 DR WPI: 94-333206/41.
 PT Recombinant DNA contg. defective gibbon ape leukaemia virus
 PT genome - and specific gene, also derived mammalian cells, virions
 PT etc., useful in gene therapy of e.g. cancer, viral infections or
 PT inherited diseases
 PS Disclosure: Page 47-52: 68pp: English.
 CC Novel replication-deficient retro virus vectors p558 (Q73733), p521
 CC (Q73734) and p537 (Q73735) include a GalV-derived packaging site,
 CC esp. the 200-910 base region of the GalV SEATO genome (Q73731), and
 CC regulatory sequences, partic. the GalV 3' long terminal repeat
 CC (Q73732). The vectors allow transfer of genes for gene therapy and
 CC transgenic animal breeding.
 SQ Sequence 10970 BP: 2595 A; 2955 C; 2847 G. 2577 T.

Query Match 42 9%; Score 2195; DB 12; Length 10970,
 Best_Local_Similarity 100.0%; Pred No 0 (00e-00);
 Matches 2195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8776 cagatggacattttcgggggaatgtgcgggaacccctattttttatttttctaaatc 8835
 QY 2903 CAGTGGGACATTTTCGGGGAATGTGCGGGAACCCCATTTTGTATTTCCTAAATAC 2962
 Db 8836 attcaaatatataccctcatgaagacaataaccctgaataatgctcaataattgaa 8895
 QY 2953 ATTCAAATGATGCGGCTCAIGAGACAATAACCCCTGATAAATGCTTCAATAATTGAA 3022
 Db 8896 aaagaaagatgatgatattcaacatttcctgtgcctcttattcccttttttgggcat 8955
 QY 3023 AAGGGAAGATGATGATGATTCACATTTTCGTGTGCGCCCTATTCCCTTTTTCGGGCAT 3082
 Db 8956 ttgccttcctgttttttgcacccagaaacgctgtgaaagtaaaagatgctgaagatc 9015
 QY 3083 TTTCCTTCCTGTTTTCCTCACCCCAACGCTGTGTAAGTAAAGATGCTGAAGATC 3142
 Db 9016 agttgggtgcagagtaggtttacatcgaactggaatctcaacacggtgaagatccttga 9075
 QY 3143 AGTTGGGTGTAAAGATGTGATGATGAACTGGAATGGAATGGAATGGAATGGAATGGA 3202
 Db 9076 gtttgcgcgcgaagaacgttttccaatgatgagcacttttaagtctgctatgcat 9135
 QY 3203 GTTTTCGCGCGGGAAGAGCTTTTCATGATGATGATGATGATGATGATGATGATGATG 3262
 Db 9136 cactattatccctgtatgacccgggcaagaaacatcgtgcgcggcggtattctc 9195
 QY 3263 CACTATTATGCGGTATTGAGGTGCGGGAAGAGTAACTGCGTGCCTGCGGTATTCTC 3322
 Db 9196 acaatgaacttgatgattactcaaccatcacaagaagaagcatttaaggatgagcagag 9255
 QY 3323 AGAATGACTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3382
 Db 9256 taaggaattatgcagtgctgccaataaccatgagtataacactgcggcgaacttacttc 9315
 QY 3383 TAAGGAATATTGAGTGTGCTGCTATACCATGATGATGATGATGATGATGATGATGATG 3442
 Db 9316 taacaacatcagaggaacccgaagagataacccttttttgcaacaacatggggatcatg 9375
 QY 3443 TCATCAACGATCGGAGACACAAAGAGATTAAGAGATTTTGTACACATGCGGGGATCATG 3502
 Db 9376 taactgccttgatcgttgggaacccggaagcgtgaatgaagccataaccaaacgacgagcgtg 9435

QY 3503 TAATCGCCTTGATCGTTGGGAACCGGAGTGAATGAAGGCAATACCAAAAGAGAGGUG 3562
 Db 9436 acaccacgagtcctgtagcaatgcacaacacactgcacaactatttaactgcgaactac 9495
 QY 3563 ACACACGATGCTGTAGCAATGCAACAAAGGTTGGTAAATATTAATGTGGGAATAC 3624
 Db 9496 ttacttagcttcccggaacaataataagactggaatgagagcgcgataaadtgcagac 9555
 QY 3623 TTACTCTAGCTTCCCGGCAACCAATTAATASATGATGAGAGGCGATAAAGTTGCAAC 3682
 Db 9556 cacttctggctggcccttcggctggctgttttatttctctgaataaattctgagccgctg 9615
 QY 3683 CACTTGTGCTGCGGACCTTGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 3742
 Db 9616 agctgggtctcccgctatcattgcaactgagccagagatgtaagccctccctatcg 9675
 QY 3743 AGCTGGGTTCGCGGTATCATTTGAGAGCTGGGCGGAGATGTAAAGGCTGCGGTATCG 3802
 Db 9676 tagttattacacgacgggagtcaggcaactatgataagcaaatagacagatcgctg 9735
 QY 3803 TAGTTATACAGGAGGAGGATGAGCAACTATGATGATGATGATGATGATGATGATGATG 3862
 Db 9736 agatagtgctcactgattaaagcattgttaactctcaacacacatttactcatatc 9795
 QY 3863 AGATAGTGCTCCTCACTGATTAAAGCATTTGTAAGCTGTAGACCAAGTTTACTCATATAC 3922
 Db 9796 tttagattgaatttaaaactcatttttaatttaaaagatctagatgaagatcccttttg 9855
 QY 3923 TTTAGATTGATTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTAAGATGCTTTTTCG 3982
 Db 9856 ataactcatgacaaaatcccttaacgtgagtttcttccactgagcgtcagaccccg 9915
 QY 3983 ATAATCTCATGACCAAAATCTTAACTGATGATGATGATGATGATGATGATGATGATG 4042
 Db 9916 taagaaagatcaagagatctcttgagatccctttttctgcgcatatctgctgctgc 9975
 QY 4043 TAGAAAGATCAAGGATCTTCTTGAATCTCTTTTTCGCGGTAAATGCTGCTTTCG 4102
 Db 9976 aaacaaaaaacacccgtaccagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 10035
 QY 4103 AACAAAAAACCCACGCTACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4162
 Db 10036 ttttccgaaggttaactggtctcagcagacgcagataaccaatactctctctctatgt 10095
 QY 4163 TTTTCCGAAGGTAACTGCTTCAAGAGAGGAGATATCAAAATATGTCCTTCTGATGCT 4222
 Db 10096 agcgtagttaggccacccttcaagaaactctgagcaccgcctacataactctgctgct 10155
 QY 4223 AGCGGTAGTTAGGCGCACTTCAAGAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 4282
 Db 10156 taatcctgttaaccagtgcgtgcagtgccgataaagctgctgtcttaccggaattggact 10215
 QY 4283 TAATCTGTTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4342
 Db 10216 caagacgatagttaccggataaaggcgccgctcagcgtggaacggggggtctgctgacac 10275
 QY 4343 CAAGACGATATTTACGCGATAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 4402
 Db 10276 agcccaacttgagagcaacacactacacgaactgaatacactacagcgtgagctatgag 10335
 QY 4403 AGCGGAGCTTGGAGCGCAAGGACCTACCTCAACTGAGTACCTACCTGAGTACCTGAG 4462
 Db 10336 aaagcgccacgctcccggaaggagaaaggcgagaggtctccgttaagcgcagcagctcg 10395
 QY 4463 AAAGCGCCAGCTTCCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4522
 Db 10396 gaacagagagcagcagagggagcttccaggggggaaacccctggtatctttatactctg 10455
 QY 4523 CAACAGGAGAGCGACAGAGGAGAGCTTCCAGGGGGAACCGCTGCTGCTGCTGCTGCTGCTG 4582
 Db 10456 tcgggttccgcaactctgaactgaagcgtcgattttttgatactcgtcagagggagcaga 10515

Cp 4712 ATGTACGCAAAAGGCGAGCAAAAGGCGAGCAAAAGGCGAGGTTGCGGCTTT 4653
Db ttcacatagactccgccccctcaacagcatcacaaaaatcaacactcaagtcacagtg 4141
Cp 4652 TTGATAGGCTGCGCTTCCCTTGAAGAGCATCAAAAATGAGCGCTCAAGTCAGAGTG 4593
Db cgaacccccacaggactataagataaccagggtttcccccctgaagctccctcgtgcgc 4201
Cp 4592 CGAAACCCGACAGGACTATAAAGATACAGAGGCTTTCCCCCTGGAAGCTCCCTCGTGGC 4533
Db tctcctgttccacccctgcgcttaccggatcacctgtccgctttctccttcgggaagc 4261
Cp 4532 TCCTCTGTTCGAGCCCTGCGGCTTACCGGATACCTGTCGCTTCTCCCTTCGCGAAGC 4473
Db gtggccttctcatagctcacgcgtgtaggtatctcagttcgggtgtaggttcggtcc 4321
Cp 4472 GTGGCGCTTTCTCATAGCTTCAGCTGTAGGTATCTCAGTTGGGTAGGTGCTTCGGTCC 4413
Db aagctgggtgtgtgcaaccccccccccttcagcccgacgctgctgccttaccggtaac 4381
Cp 4412 AAGCTGGGTGTGTGCACGAAGCCCGCTTCAGCCGAGCGCTGCGCTTATCCGGTAAAC 4353
Db taccgtcttgactcaaacccggtgaagacacgacttatcgccactggcagcagcaactcgt 4441
Cp 4352 TATCGCTTTGAGTCCAAACCCGCTTAAAGACACGACTTATGCGCACTGGCAGCCACCTGGT 4293
Db aacagattagcagagcagggtatgtaggcggtgctacagagttcttgaagtgtgcct 4501
Cp 4292 AACAGGATTAGCAGAGCGAGGTATGTAAGCGGTGTCTACAGAGTCTTGAAGTGTGGCT 4233
Db aactacggttacactagaaggacagatattgtgtatctcgtctgctgaagccagttacc 4561
Cp 4232 AAGTAGGGTACACTASAGGACAGTATTGTGTAATCTGCTGCTGCTGAAGCCAGTTACC 4173
Db ttgcgaaaagagttgtagctcttgatccgcgcaacaaacacacgctgtagcggagt 4621
Cp 4172 TTGCGAAAAGAGTTGTAAGTCTTGTATCCGCAACCAACACCCGCTGGTAGCGGTGGT 4113
Db tttttgtttgcaagcagcagatattcgcgcagaaaaaaggatctcaagaagatcccttg 4681
Cp 4112 TTTTITTTTGAAGTACGATTAAGGACAGTATAGGCGAGCAAAAGGATCTCAAGAAGATCCITG 4053
Db atcttttctacgggtctgaagctcaagtcagtcagcaaaactcaagtcagtcagtcagtc 4741
Cp 4052 ATCTTTCTAAGGGGCTTCAAGCTCAGTGTGAAGCAAGAACTCAGGTAAAGGATTTGGTC 3993
Db atcagattatcaaaaaggatcttcacctagatcccttttaaaatcaaaatgaagttttaa 4801
Cp 3992 ATCAGATTATCAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTANA 3933
Db tcaatcaaatatataatagtaaaacttgctgacagttaccaatgcttaactcagtgag 4861
Cp 3932 TCAATCTAAATATATAGTAAGTAACTTGTGTGACGTTACCAATGCTTAAATCAGTGAG 3873
Db gacccatctcagcagctctgcttatcttctcctatccatagttgctgactcccgctcgtg 4921
Cp 3872 GCACCTATCTCAGCGATCTGTATTCTGTTATCTATCTATCTATCTATCTATCTATCTATCT 3813
Db tagataactacgatacgggaggggttaaccatctgcccccauigtgcaatgaataaccgga 4981
Cp 3812 TAGATAACTATAGTATAGGAGAGGCTTACATATCTGAGTCTGCTGCTGCTGCTGCTGCTG 3753
Db gacccagctcaccgctctcagatttatcagcaataaaacccagccgagcggagcggag 5041
Cp 3752 GATCCACGCTCAGCGCTCTCAGATTATCAGCAATAAAACAGCCAGCGGAGGCGGAG 3693
Db cagagaagtgctcctgcaactttatccgctccatccagctctatttaattgttgcgggaa 5101
Cp 3692 CCAAGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATGTGTGGGAA 3633
Db qctagaatgaagttcgccagtttaagtttgcgcaacggttggccattgctacaggc 5161
Cp 3632 GCTAATAAIAAGTATAGGAGTATAGTATAGGAGTATAGGAGTATAGGAGTATAGGAGTATAGGAGT 3573

RESULT 10

ID Q39050 standard: DNA: 6824 BP.
AC Q39050:
DT 28-JUL-1993 (first entry)
DE K.lactis/S. cerevisiae genetic vector.
KW Genetic vector; integration; Kluyveromyces lactis; 25S ribosomal DNA;
KW Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;
KW Expression cassette; HIS3; marker; transformant; human; lysozyme; HL2;
KW GAL7; signal sequence; killer toxin; transcription termination signal;
KW FLP; 2 micron plasmid; ss.
CS Synthetic.
PN EP-537456-A.
PD 21-APR-1993.
PF 31-AUG-1992: 114838.
PR 04-SEP-1991: IT-M12349.
PA (ISTS) SCLAVO SPA.
PI Galeotti CU, Gallo E, Riccio ML, Rossolini GM, Thaller MC;
DR WPI: 93-127394/16.
PT Vector for Kluyveromyces lactis and Saccharomyces cerevisiae -
PT which allows stable multiple integration of DNA for prodn. of
PS heterologous proteins
PS Claim 1: Fig 1: 26pp. English.
CC This sequence represents a genetic vector which allows the stable


```

QY 4320 GTGTCCTTATCGGGTGTGAACTAAGACGATAGTATACGGGATAAGGCGACAGCGTGGG 4379
Db 6026 CTGAACGGGGGTTCTGTCACACAGCCAGCTGTGAGCAAGCACTACACCGAAGTGA 6085
QY 4380 CTGAACGGGGGTTCTGTCACACAGCCAGCTGTGAGCAAGCACTACACCGAAGTGA 4439
Db 6086 ATACCTACAGCTGAGCTATGAGAAGCGCCACGCTTCCGGAAGGAGGAGCGGACAG 6145
QY 4440 ATACCTACAGCTGAGCTATGAGAAGCGCCACGCTTCCGGAAGGAGGAGCGGACAG 4499
Db 6146 GTATCCGTAAGCGGCGAGGTCGGAACAGAGAGAGCGAGCGAGGAGGAGGAGGAGG 6205
QY 4500 GTATCCGTAAGCGGCGAGGTCGGAACAGAGAGAGCGAGCGAGGAGGAGGAGGAGG 4559
Db 6206 CAGCTGATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6265
QY 4560 CAGCTGATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4619
Db 6266 GTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6325
QY 4620 GTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4679
Db 6326 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6385
QY 4680 GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4739
Db 6386 TGTGATAACCGTATTACCGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6445
QY 4740 TGTGATAACCGTATTACCGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4799
Db 6446 CAGAGCGAGCGAGTGTGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6505
QY 4800 CAGAGCGAGCGAGTGTGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4859
Db 6506 CCGCGCGCTTGGCGCTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 6565
QY 4860 CCGCGCGCTTGGCGCTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4919
Db 6566 GAGCGAGCGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 6625
QY 4920 GAGCGAGCGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4979
Db 6626 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6685
QY 4980 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5039
Db 6686 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6718
QY 5040 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5072

RESULT 11
ID V02042 standard. DNA. 7287 BP.
AC V02042;
DI 08-JUN-1998 (first entry)
DE Plasmid pWRG3169 encoding murine interleukin-12.
KW interleukin-12; IL-12; cytokine; growth factor; mouse;
KW plasmid pWRG3169; cancer; tumour; metastasis; gene therapy; ds:
KW cyclic circular
OS Chimeric - Mus musculus.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Bos taurus.
OS Chimeric - Rhesus macaque polyoma virus.
FH Key Location/Qualifiers
FT promoter 1..628
FT /tag= a
FT /note= "CMV promoter"
FT iDNA 629..810
FT /tag= b
FT CDS 953..1673
FT /tag= c
FT /product= p35 subunit
FT /note= "contains an intron"

```

```

FT intron 1259..1331
FT /tag= d
FT polyA_site 1797..2024
FT /tag= e
FT /note= "Bovine growth hormone polyA site"
FT promoter 2110..2737
FT /tag= f
FT /note= "CMV promoter"
FT iDNA 2738..2919
FT /tag= g
FT CDS 2983..3990
FT /tag= h
FT /product= p40 subunit
FT polyA_site 4075..4306
FT /tag= i
FT /note= "Bovine growth hormone polyA site"
FN W09746263-A1.
PD 11-DEC-1997.
PF 04-JUN-1997; U09591.
PR 05-JUN-1996; US-659206.
PA (AURA) AURAGEN INC.
PI Rakhmillevich AL Yang N;
DR WPI: 98-041898/04.
DR P-PSDB: W44004-05.
PT Interleukin-12 gene therapy of tumours - comprises delivering
PT construct containing promoter and sequences encoding interleukin-12
PT p35 and p40 subunit(s) to target cells in vivo
PS Claim 5; page 24-30; 50pp; English.
CC Plasmid pWRG3169 is a tandem plasmid encoding both subunits, i.e.
CC p35 (see W44004) and p40 (see W44005), of murine interleukin-12
CC (IL-12). Each subunit gene was cloned from a mouse spleen cDNA
CC library and has been placed under the transcriptional control of
CC a separate cytomegalovirus (CMV) promoter. An SV40 splicing
CC donor/splicing acceptor is provided between each subunit gene and
CC its CMV promoter. The backbone of the plasmid is from pUC19.
CC pWRG3139 induces at least twice the expression of IL-12 as the
CC bicistronic vector pWRG3136 (see V02043) in vivo and in vitro. A
CC novel method of treating tumours in a mammal involves delivering
CC copies of an expressible foreign genetic construct, especially
CC pWRG3169 or pWRG3136, comprising a promoter operative in the
CC mammalian epidermal cells and DNA sequences encoding p35 and p40
CC subunits of IL-12 to target cells in vivo. Delivery of the
CC construct allows IL-12 expression for treatment of solid,
CC metastatic or disseminated tumours, and regression of established
CC tumours. The treatment is effective even when the genetic
CC construct is delivered to a site distant from the tumour.
SQ Sequence 7287 BP; 1798 A; 1877 C; 1845 G; 1766 T;

```

```

Query Match 41.9%, Score 2141, DB 39, Length 7287;
Best Local Similarity 99.4%; Pred. NO 0.00e+00;
Matches 2155; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 5004 aggtggacacttttcgggaaatgacgagcaacccctattttatttttctaaataca 5063
QY 2904 AGTGGCACTTTTCGGGAAATGTGCGGAAACCCCTATTGTTTATTTTCTAAATACA 2963
Db 5064 ttcaaatatgtatccgctcatgagacaataaacctgtataaatgcttcaataattgaaa 5123
QY 2954 TTCAAATATGTATCCGTCATGAGACATAAACCTGTATAATGCTTCAATATATGAAA 3023
Db 5124 agggaagagatgagatattcaacatttcctgtgcgccttattcccttttttgcgcatt 5183
QY 3024 AAGGAAGAGATGATGATATCAACATTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3083
Db 5184 ttgccttcctgtttttgtctcaccagaaacgctgtgtaaaataaaatgctgaagatca 5243
QY 3084 TTGCTTCTGCTGTTTGTCTCACCAGAGAAACGCTGTGTGAAAGTATGCTGAGATCA 3143
Db 5244 gttgggtgcacgagtggtttacatcgaaactggtatctcacacgagtgatccttgag 5303
QY 3144 GTTGGGTGCACGAGTGGGTATCATCGAACTGGATCTCAACAGCGGTAAGATCTCTGAG 3203
Db 5304 ttttcgcccgaagaacgttttccaatgatgagcaacttttaagtctgtatgtgcgcgc 5363

```



```
Cp 4052 ATCTTTCTACGGGCTCTCAAGCTCAGTGTGAAGAAACACACGTTAAGGGAATTTGGTC 3993
Db 8894 atcaaatatacaaaaggatcttcacccatagatcccttttaaatataaaatgaatttttaaa 8953
Cp 3992 ATGAGATTATCAAAAGGATCTTACCTAGATCCCTTTTAAATTAATAATGAAGTTTATAA 3933
Db 8954 tcaatcaaatatataatgaataaacttggtctgcagcttaccatgcttaataatcagtgag 9013
Cp 3932 TCAATCTAAAGTATATAGTAACCTTGGCTGACAGTTACCAATGCTTAATCAGTGAG 3873
Db 9014 gacattatcagagatctctatttctgcttcattcattcagcttgcctgaactcccccgtgtg 9073
Cp 3872 GCACCTATCTCAGGATCTCTATTCTGTTGCTATCATCATAGTTGCTGACTCCCGCTGTG 3813
Db 9074 tagataactacatagcagagagcttaccatctgcgccacgtctgcaatgataaccgcga 9133
Cp 3812 TATATAAATTCGATACGGAGGGGCTTACCACTGAGGCTGCTGCTGCTGCTGCTGCTG 3753
Db 9134 gacccacgtccacccgctccagatttatcagcaataaaaccagccagcgggaagggccgag 9193
Cp 3752 GACCCAGGCTCAGCGCTCCAGATTATCAGCAATAAACACGACGCGGGAAGGCCGAG 3693
Db 9194 cagagaagtggctcgtgcaatttatccgcctccatccagcttataattgttgcgggaa 9253
Cp 3692 CGAGAAAGTGGTCTGCAACCTTATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3633
Db 9254 gctagagtaagttagtgcgcagcttaagtgttgcgaacgtgttgccttgcctgctacagc 9313
Cp 3632 GTAGAGTAAGTASITGCGCAATTAATAGTATTCGCAACAGTGTGCTGCTGCTGCTGCTG 3573
Db 9314 atgtggtatcagctgtgctgttggatggattcattcagctcgcgttcccaacgatca 9373
Cp 3572 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3513
Db 9374 agcgaattacatgatacccccagcttgcgaataaaaggggttagctccttcggtccctcg 9433
Cp 3512 AGCGAGTATACATGATCCCGCAATGTTGCAAAAAGGGTGTAGCTCTTCGCTCCCTCCG 3453
Db 9434 atgttgcagagtaagttagtgcgcagctgttatcactcagctgttgcctgagcagctcgt 9493
Cp 3452 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3393
Db 9494 aattccttactgctcagctcagctcagctcagctcagctcagctcagctcagctcagctcag 9553
Cp 3392 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3333
Db 9554 aagctctctgagaatagtgatgcgcagcagcttgccttgccttgccttgccttgccttgcct 9613
Cp 3332 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3273
Db 9614 gataatccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 9673
Cp 3272 GATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3213
Db 9674 gacgcaaaactctcagagcttaccgctgttgagatccagctcagctcagctcagctcagctcag 9733
Cp 3212 GCGGCAAAATCTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3153
Db 9734 gacgcaaaactctcagagcttaccgctgttgagatccagctcagctcagctcagctcagctcag 9793
Cp 3152 GCAGCTCAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3093
Db 9794 ggaagcaaaatgcccaaaaggaataaggggagacagcagaaatgttgaatactcata 9853
Cp 3092 GGAAGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3033
Db 9854 ctcttctcttttgaatattatgaagcatttatcagaggttatgtctcattcagcagagatac 9913
Cp 3032 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2973
Db 9914 atatttgaatgtatttagaaaaataaacaataggggttccgcgcacatttcccccgaata 9973
Cp 2972 ATATTGAAATGATTTAGAAAAATAAACCAATAGGGGTTCCCGGCAATATTCCTCCGGA 2913
```

```
Db 9974 gtgccacctg 9983
Cp 2912 GTGCCACCTG 2903

RESULT 14
ID Q40279 standard; DNA: 4145 BP.
AC Q40279;
DE 02-AUG-1993 (first entry)
KW Sequence of clone pS2gpt-S4.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..2226
FT /tag= a
FT /label= pN2gpt-S4
FT /note= "Posn. 1 corresp. to the first nucleotide
FT G '5'-TGGCAGCTTT TCGGGGAAAT-3'."
FT misc_feature 2227..2236
FT /tag= b
FT /label= SmaI adaptor
FT 2396..2851
FT /tag= c
FT /label= E. coli gpt gene
FT complement (3081..3323)
FT /tag= d
FT /label= vaccinia P7.5 promoter
FT 3358..3451
FT /tag= e
FT /label= S4 of pN2gpt-S4
FT /note= "oligo P-artp(9)."
FT misc_feature 2237..4145
FT /tag= f
FT /label= pN2gpt-S4

AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; C21269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;
WFI; 93-156461/16.
DR Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example; Pages 160-162; 206pp; English.
CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. The plasmid
CC pN2gpt-S4 was digested with XbaI and ligated with a SmaI-adaptor
CC inactivating the XbaI and creating a SmaI site. The resulting
CC plasmid was designated pS2gpt-S4. The S4-promoter segment of
CC lasmid pS2gpt-S4 was removed by cleavage with pStI and HpaI and
CC replaced with a 172 bp pStI-HpaI P2-promoter segment. The final
CC plasmid was designated pS2gpt-P2.
SQ Sequence 4145 BP; 1067 A, 1007 C, 1039 G, 1032 T,
Query Match 41.8%; Score 2138; DB 7; Length 4145;
Best Local Similarity 99.4%; Pred. NO 0.00e+00;
Matches 2152, Conservative 0; Mismatches 14, Indels 0, Gaps 0;

Db 1 gtgccactttccggagaaatgtgcgagaaacccctatttttttttaataacatt 60
QY 2906 GTGCCACTTTTCGGGGAATGTGCGGGAAGCCCTTATTTGTTTATTTCTAAATACAT 2945
Db 61 caaatatgtatccgtcatgacacaataaccctgataaactcctcaatattgaaaaa 120
QY 2965 CAAATATGTATCCGCTCATGAGACAATAACCTGTAATAATGCTTCATATATATT 3025
Db 121 gaaagagatgagttattcaaacatttccatgtgcgccttattcccttttttgcggcattt 180
```


OS Chimaeric Homo sapiens:
 OS Chimaeric synthetic:
 FH Key
 FT 2218..4335
 FT cds /tag- a
 PN W09725345.A1.
 PD 17-JUL-1997.
 PF 03-JAN-1997: U00140.
 PR 05-JAN-1996: US-009629.
 PA (ELIL) LILLY & CO ELI.
 PI Kovacevic S, Otto KA, Rao RN:
 DR WPI: 97-372814/34.
 DR P-PSDB: W18573.
 PI Fusion protein comprising human cyclin and cyclin-dependent protein
 PT kinase - useful for research in cell cycle regulation allowing
 PT addition of both components simultaneously, giving greater control
 PT over reaction conditions
 PS Disclosure: Page 17-22: 57pp. English.
 CC A DNA sequence (J69891) in plasmid K480, deposited in E. coli as
 CC NPPI R-21491, includes a coding region for a fusion protein
 CC (W18573) comprising human cyclin D1 linked to human CDK4, flanked
 CC by N-terminal histidine residues, a myc epitope and a thrombin
 CC cleavage site, and by a C-terminal streptavidin binding domain, to
 CC facilitate purification. The fusion protein can be expressed in
 CC host cells, pref. using a baculovirus expression system, for use in
 CC research into cell cycle regulation.
 SQ Sequence 4540 BP: 1067 A: 1180 C: 1253 G: 1040 T:
 Query Match 41.8% Score 2138: DB 32: Length 4540:
 Best Local Similarity 99.7%: Prod No 0 00e-00:
 Matches 2154: Conservative 0: Mismatches 16: Indels 0: Gaps 0:
 Db 70 cagtgagcaactttcggggaataatgcgggaacccctatttggttatttttctaaataac 129
 QY 2503 CAGSISGACATTTTGGGSSAAATG180GSSAAACCCATTTTGTATTTCIACAAATAC 2962
 Db 130 attcaaatatcatcgctcatgaagcaataaccctatgaataatgcttcaataattgaa 189
 QY 2563 ATTCAAA1A1S1ATGCGTCTATGAGAAATAA122TGATAAAATGCTTCAATAATTGAA 3022
 Db 190 aaagaaagatagatattcaacatttcggtgcgccttattcccttttttggcgcat 249
 QY 3023 AAAGTAAATATGAGTATTCACATTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGG 3082
 Db 250 ttgccttctggttttgcacccagaaacgctggtgaaatgaaatgctgaagatc 309
 QY 3083 TTTCGCTTCCTTTTTCGTCACCAAGAAAGCTGTAAGATGCTGAAGATC 3142
 Db 310 agttggatgcaagagtggtttacatgaactggatgctgaagagggatgcttgaga 369
 QY 3143 AGTTGGTGTGCACAGTGGGTTACATGAACTGATCAGAGCTGAAAGCTGCTTGAGA 3202
 Db 370 gtttgcgcgcgaagaggttttccatgataggaactttttaaagtctctgtatggcg 429
 QY 3203 GTTTTCGCGCGGAAAGACGTTTTCGAATGATGAGCACTTTTAAAGTCTGCTATGCA 3262
 Db 430 cgttattatccctattgacgcggggaagagcaactcgtgcgcgcgcgcgcgcgcgcgc 489
 QY 3263 CACTATTATCCGCTATTGACGCGCGGGAAGAGTAACTCGCTGCGCGCGCGGTATTCTC 3322
 Db 490 aqaatgactggttgatgactaccagtcacagaaagacatcttacggatggcatgaacag 549
 QY 3323 AGAATGACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3382
 Db 550 taagaaattatcagtgctgacataaccatgagtataacactgcggcgcaacttacttc 609
 QY 3383 TAAAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3442
 Db 510 tcaacaacatcgagaggaacaaagagactaacgccttttttgcacaacatgggggatc 669
 QY 3443 TGAAAGCAATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3502
 Db 570 taactcgccttgatgctggtggaacccgagctgaatgaagaaaccataccaaagcagag 729

QY 3503 TAACTCGCTTATGCTTGGGAAACGGAGCTTGAATGAAGGCAATGAATGAAGGCAAGGCTG 3562
 Db 730 acacaacgatgctgtagcaatggcaacaacggttgcgcaaacatttaactggcgaactac 789
 QY 3563 AAGACGSAIGCTGTAGCAATGGCAACAAAGTTGGGCAAACTATTAACGTGGCAATAC 3622
 Db 790 ttacttaacttcccgccgcaacaattaatagactgagatggagcggaataaagtgtcagac 849
 QY 3623 TTACTCTAGCTTCCCGGGAATTAATTAATAGACTGATGAGAGGAGGAGGAGGAGGAG 3682
 Db 850 cacttctgctgctggcccttccggctggctggtttattgtgtgataaactctgagaccggtg 909
 QY 3683 CACTTCTGAGCTGCTGCTTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 3742
 Db 910 agcgtgctgctgctggctgctgctgctgctgctgctgctgctgctgctgctgctgctg 969
 QY 3743 AGCGTGGGCTTGGGCTGATCATTTGAGACTGGGCTGAGAGTGGGCTGAGAGTGGGCT 3802
 Db 970 tagttatctacacgagcggggagctgagcaactatggatgagcaagaaatagacatatctg 1029
 QY 3803 TAGTTATCTACAGACGGGGAGTCAAGCACTATGGATGAGCAATAGAGATAGAGATCG 3862
 Db 1030 aatatggtcctcactgattgaagcatttgtaactctcagacccaacttactcatataac 1089
 QY 3863 AGATAGTGCCTCACTGATTAAAGCATTTGATGCTGAGCACTGAGCACTGAGCACTGAG 3922
 Db 1090 tttaaatgatttaaaacttattttaaataaaagcatctagctgaagactccttttgc 1149
 QY 3923 TTAGATGATTTAAACTTCATTTTAAATTAAGGATCTAGGTAAGATCTCTTTTTC 3982
 Db 1150 ataactcctgaccaaaatcccttaactgaggttttctccactgagcgtcagaccccg 1209
 QY 3983 ATAACTCATGACCAAAATCCCTTAACGTGAGTTTCCTTCCACTGAGCGTCAAGCCGG 4042
 Db 1210 tagaaagatcaaaagatcttcttgagatccttttttctgctgctgctgctgctgctgct 1269
 QY 4043 TAGAAAGATCAAGAGATCTTCTTGAATCTCTTTTCTGAGGTAATCTGAGGTAATCT 4102
 Db 1270 aaacaaaaaaccccgctaccagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1329
 QY 4103 AAACAAAAAACCCGCTACAGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4162
 Db 1330 ttttccgaggttaactggttcacagcagcagcagcagcagcagcagcagcagcagcagc 1389
 QY 4163 TTTTCCGAAGGTAACTGCTTTCAGCAGAGCGCAGATACCAAACTAGTCTGCTTCTAGTGT 4222
 Db 1390 agcgtgagtaggacacacttcaagaactctgtagaccccgctacataactcgtctctgc 1449
 QY 4223 AGCGTAGTTAGCGCACCACTTCAAGAACTCTGTAGGACCGGCTACATACTGCTGCTG 4282
 Db 1450 taactcgttaccagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1504
 QY 4283 TAATCTGTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4342
 Db 1510 caagcagtagttaccgagataaggcgcgcgctggcggtgacgggggggttctgctgacac 1569
 QY 4343 CAAGAGATAGTTACCGGATAAGGGGCAAGCGCTGGGCTGAGGAGGAGGAGGAGGAGGAG 4402
 Db 1570 agccagcttgagaggaacacactacacgaactgaataactacagcgtgaagcattgag 1629
 QY 4403 AGCGGAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4462
 Db 1630 aaagcccccgcgttcccgagggagaaaggcgagcaggtatccggttaagcgcgagggctg 1689
 QY 4463 AAAGCGCGACGCTTCCCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4522
 Db 1690 gaacaggagagcgacgggggagcttccagggggaaacgccttggtatctttatagtcgtg 1749
 QY 4523 GAACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4582
 Db 1750 tggggttccgcacactctgacttgagcgtcgtatttttctgactcgtcagagggagga 1804

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrchann n.d. - n.d. database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 14:33:11 1999; MasPar time 6493.81 Seconds
Tabular output not generated. 1410.547 Million cell updates/sec

Title: >US-09-020-716-5
Description: (1-5115) from US09020716.seq
Perfect Score: 5115
N.A. Sequence: 1 GTTGGAGGCTCTCCCATATGTCAAGCTATGATCCCAACGC 5115
Comp: CAACCTCTGAGAGGCTATACAGTGTGATAGCTAGGTGGC

Scoring table: TABLE default
Gap 5

Nmatch STD : Dbase 0: Query 0

Searched: 2275926 seqs, 895388244 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database: emb1-est56
Database: 1:em_est1 2:cm_gss1 3:em_gss2 4:em_gss3
genbank-est109

5:qb_est1 6:qb_est10 7:qb_est11 8:qb_est12 9:qb_est13
10:qb_est14 11:qb_est15 12:qb_est16 13:qb_est17
14:qb_est18 15:qb_est19 16:qb_est2 17:qb_est20
18:qb_est21 19:qb_est3 20:qb_est4 21:qb_est5 22:qb_est6
23:qb_est7 24:qb_est8 25:qb_est9 26:qb_gss1 27:qb_gss2
28:qb_gss3 29:qb_gss4

Statistics: Mean 12.493; Variance 1.996; scale 6.260

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	555	12.8	721	26	AG004662 Homo sapiens genomic D	0.00e+00
2	555	12.8	721	27	AG011001 Homo sapiens genomic D	0.00e+00
3	538	12.5	711	27	AG010947 Homo sapiens genomic D	0.00e+00
4	538	12.5	711	26	AG004508 Homo sapiens genomic D	0.00e+00
5	533	12.4	718	26	AG004363 Homo sapiens genomic D	0.00e+00
6	533	12.4	718	27	AG010489 Homo sapiens genomic D	0.00e+00
7	530	12.3	692	27	AG040352 CIT-HSP-2327K21.TF CIT	0.00e+00
8	528	12.3	698	27	AG009976 Homo sapiens genomic D	0.00e+00
9	528	12.3	698	26	AG003787 Homo sapiens genomic D	0.00e+00
10	624	12.2	720	28	AG013858 Homo sapiens genomic D	0.00e+00
11	624	12.2	720	26	AG006062 Homo sapiens genomic D	0.00e+00
12	617	12.1	652	28	AG010863 CIT-HSP-2379M3.TF CIT	0.00e+00
13	619	12.1	747	25	AG007052 Homo sapiens genomic D	0.00e+00

C 14	612	12.0	666	27	AQ079096	CIT-HSP-2356E24.TF CIT	0.00e+00
C 15	609	11.9	640	27	AQ074298	CIT-HSP-2382L21.TF CIT	0.00e+00
C 16	598	11.7	696	27	AG009765	Homo sapiens genomic D	0.00e+00
C 17	598	11.7	696	26	AG003576	Homo sapiens genomic D	0.00e+00
C 18	596	11.7	748	17	A1124281	1 5a thymophila 8-12 h	0.00e+00
C 19	591	11.6	542	28	AQ13213	CIT-HSP-2372L22.TF CIT	0.00e+00
C 20	592	11.6	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C 21	595	11.6	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C 22	581	11.4	594	26	C1786	Ciona intestinalis ger.	0.00e+00
C 23	579	11.3	634	27	AQ38010	CIT-HSP-2328L2.TF CIT	0.00e+00
C 24	578	11.3	653	27	AQ37504	CIT-HSP-2365K21.TF CIT	0.00e+00
C 25	577	11.3	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C 26	573	11.2	633	27	AQ04787	CIT-HSP-2336K22.TF CIT	0.00e+00
C 27	574	11.2	661	27	AQ057603	CIT-HSP-2345M3.TF CIT	0.00e+00
C 28	569	11.1	637	27	AQ077360	CIT-HSP-2365E81.TF CIT	0.00e+00
C 29	564	11.0	615	28	AQ112424	CIT-HSP-2376E2.TF CIT	0.00e+00
C 30	563	11.0	625	28	AQ111342	CIT-HSP-2371M3.TF CIT	0.00e+00
C 31	559	10.9	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C 32	547	10.7	600	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00
C 33	545	10.7	623	27	AQ081172	CIT-HSP-2357C31.TF CIT	0.00e+00
C 34	540	10.6	612	27	AQ080586	CIT-HSP-2382D21.TF CIT	0.00e+00
C 35	530	10.4	605	27	AQ077353	CIT-HSP-2365C21.TF CIT	0.00e+00
C 36	533	10.4	682	27	AQ074593	CIT-HSP-2361L23.TF CIT	0.00e+00
C 37	526	10.3	587	27	AQ079166	CIT-HSP-2382M2.TF CIT	0.00e+00
C 38	526	10.3	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C 39	525	10.3	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C 40	525	10.3	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C 41	528	10.3	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C 42	524	10.2	540	27	AQ074307	CIT-HSP-2382N1.TF CIT	0.00e+00
C 43	521	10.2	597	27	AQ040843	CIT-HSP-2324K24.TF CIT	0.00e+00
C 44	510	10.0	664	26	AG005286	Homo sapiens genomic D	0.00e+00
C 45	510	10.0	664	27	AG012931	Homo sapiens genomic D	0.00e+00

ALIGNMENTS

RESULT LOCUS	1	AG004662	721 bp	DNA	GSS	30-JAN-1998
DEFINITION		Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic survey sequence.				
ACCESSION		AG004662				
NID		g2826191				
KEYWORDS		GSS.				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens DNA, clone: 879G12X91.				
REFERENCE		1 (bases 1 to 721)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Homo sapiens genomic DNA, chromosome 21q				
JOURNAL		Published Only in Database (1998) In press				
REFERENCE		2 (bases 1 to 721)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.				
		Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@bc.ims.u-tokyo.ac.jp. Tel:0427-78-9732. Fax:0427-78-9561)				

FEATURES

FEATURES	Location/Qualifiers
source	1..721
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="21"
	/clone="879G12X91"
	/map="21q"
BASE COUNT	195 a 163 c 175 g 179 t 9 others
ORIGIN	
Query Match	12.8%, Score 655, Db 26, Length 721;
Best Local Similarity	94.6%; Pred No 0.00e+00;
Matches	682, Conservative 0, Mismatches 6, indels 4, gaps 4;

REFERENCE 1 (bases 1 to 711)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 711)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (28-JUL-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
 (E-mail:hattori@shgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
 Fax:0427-78-9561)
 Location/Qualifiers
 1..711
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="S594BG38"
 /map="21q"
 BASE COUNT 169 a 188 c 172 g 182 t
 ORIGIN
 Query Match 12.5% Score 638, DB 27, Length 711
 Best Local Similarity 97.7%, Pred. No 0.00e+00,
 Matches 669; Conservative 0; Mismatches 13; Indels 3; Gaps 3;
 Db 28 TAGATAACTACCATCAGGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCG 87
 Cp 3812 TAGATAACTACCAT-ACGGAGGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCG 3754
 Db 88 AGACCCAGGCTCAGCGGCTCCAGATTTATAGCAATAAATCAAGCAAGCGGAGAGGCGGA 147
 Cp 3753 AGACCCAGGCTCAGCGGCTCCAGATTTATAGCAATAAATCAAGCAAGCGGAGAGGCGGA 3694
 Db 148 GGGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTATTTGTCGCGGGA 207
 Cp 3693 GGGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTATTTGTCGCGGGA 3634
 Db 208 AGCTAGAGTAAAGTGTGGCGGAGTTATAGTTTGGCAAGGTTGTGCCATTGCTACAGG 267
 Cp 3633 AGCTAGAGTAAAGTGTGGCGGAGTTATAGTTTGGCAAGGTTGTGCCATTGCTACAGG 3574
 Db 268 CATCGTGTGTGACGCTGGTGTGGTATGGTTCATTCAGTCTCCGCTTCCCAACGATC 327
 Cp 3573 CATCGTGTGTGACGCTGGTGTGGTATGGTTCATTCAGTCTCCGCTTCCCAACGATC 3514
 Db 328 AAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCTCTCC 387
 Cp 3513 AAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCTCTCC 3454
 Db 388 GATCGTTGTGAGAAGTAAAGTTGGCGGAGTGTATGCTATGCTATGCTATGCTACAGG 447
 Cp 3453 GATCGTTGTGAGAAGTAAAGTTGGCGGAGTGTATGCTATGCTATGCTACAGG 3394
 Db 448 TAATTCCTTACTGTGATGCGATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAAC 507
 Cp 3393 TAATTCCTTACTGTGATGCGATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAAC 3334
 Db 508 CAAGTCATCTCTGAGAAATAGTGTATGGGCGCACTGAGTGTGCTTGGCGGCGTCAATACG 567
 Cp 3333 CAAGTCATCTCTGAGAAATAGTGTATGGGCGCACTGAGTGTGCTTGGCGGCGTCAATACG 3274
 Db 569 GGATAATACGGCGCCACATAGCAGAACTTTAAAGTGTGCTATCATTTGGAAGCTTCTTC 627
 Cp 3273 GGATAATAGTGTATGATATGAGCAAGTCTTAAAGTGTGCTATCATTTGGAAGCTTCTTC 3214
 Db 628 GGGGCGCAAACTCTCAGAGATCT-ACCGTGTGTGAGATCTCAGTTCGATGTAAACCCACTC 686
 Cp 3213 GGGGCGCAAACTCTCAGAGATCT-ACCGTGTGTGAGATCTCAGTTCGATGTAAACCCACTC 3155
 Db 687 GTGCACCCCAACTGATCTTCCAGCATC 711
 Cp 3154 GTGCACCCCAACTGATCTTCCAGCATC 3130

4 AG004608 711 bp DNA GSS 29-JAN-1998
 LOCUS Homo sapiens genomic DNA, 21q region, clone, S594BG38, genomic
 DEFINITION survey sequence.
 ACCESSION AG004608
 NID 92822058
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone:S594BG38.
 ORGANISM Homo sapiens
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 711)
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 711)
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
 (E-mail:hattori@shgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
 Fax:0427-78-9561)
 Location/Qualifiers
 1..711
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="S594BG38"
 /map="21q"
 BASE COUNT 169 a 188 c 172 g 182 t
 ORIGIN

Query Match 12.5% Score 638, DB 26, Length 711
 Best Local Similarity 97.7%, Pred. No 0.00e+00,
 Matches 669; Conservative 0; Mismatches 13; Indels 3; Gaps 3;
 Db 28 TAGATAACTACCATCAGGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCG 87
 Cp 3812 TAGATAACTACCAT-ACGGAGGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATACCGCG 3754
 Db 88 AGACCCAGGCTCAGCGGCTCCAGATTTATAGCAATAAATCAAGCAAGCGGAGAGGCGGA 147
 Cp 3753 AGACCCAGGCTCAGCGGCTCCAGATTTATAGCAATAAATCAAGCAAGCGGAGAGGCGGA 3694
 Db 148 GGGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTATTTGTCGCGGGA 207
 Cp 3693 GGGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTATTTGTCGCGGGA 3634
 Db 208 AGCTAGAGTAAAGTGTGGCGGAGTTATAGTTTGGCAAGGTTGTGCCATTGCTACAGG 267
 Cp 3633 AGCTAGAGTAAAGTGTGGCGGAGTTATAGTTTGGCAAGGTTGTGCCATTGCTACAGG 3574
 Db 268 CATCGTGTGTGACGCTGGTGTGGTATGGTTCATTCAGTCTCCGCTTCCCAACGATC 327
 Cp 3573 CATCGTGTGTGACGCTGGTGTGGTATGGTTCATTCAGTCTCCGCTTCCCAACGATC 3514
 Db 328 AAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCTCTCC 387
 Cp 3513 AAGGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTCTGGTCTCTCC 3454
 Db 388 GATCGTTGTGAGAAGTAAAGTTGGCGGAGTGTATGCTATGCTATGCTATGCTACAGG 447
 Cp 3453 GATCGTTGTGAGAAGTAAAGTTGGCGGAGTGTATGCTATGCTATGCTACAGG 3394
 Db 448 TAATTCCTTACTGTGATGCGATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAAC 507
 Cp 3393 TAATTCCTTACTGTGATGCGATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAAC 3334
 Db 508 CAAGTCATCTCTGAGAAATAGTGTATGGGCGCACTGAGTGTGCTTGGCGGCGTCAATACG 567
 Cp 3154 CAAGTCATCTCTGAGAAATAGTGTATGGGCGCACTGAGTGTGCTTGGCGGCGTCAATACG 3130


```

Db 570 CUAAGTTATCTG-TATAACATTTAGATTGATTTAAACTTTCATTTTAAATTTAAAGGAT 628
|||||
Qy 3903 CCAAGTTTACTATATATCTTTAGATTGATTTAAACTTTCATTTTAAATTTAAAGGAT 3962
|||||
Db 629 CTAGGTCAGATUNTTT-GATAATCTCATGACCAAAATTCCTTAACTGAGTTTCGTT 687
|||||
Qy 3963 CTAGGTCAGATUNTTT-GATAATCTCATGACCAAAATTCCTTAACTGAGTTTCGTT 4022
|||||
Db 688 TCCACTGA 695
|||||
Qy 4023 -CCACTGA 4029
|||||

RESULT 10 AG013858 720 bp DNA GSS 10-SEP-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic
DEFINITION survey sequence.
ACCESSION AG013858
NID g3560348
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 762015SpN14.
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published only in Database (1998) In press
REFERENCE 2 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1998) to the DBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228 Japan
(E-mail:hattori@ngs.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)

FEATURES
source
1 720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="762015SpN14"
/map="21q"
BASE COUNT 172 a 178 c 170 g 186 t 14 others
ORIGIN
Query Match 12.2%, Score 624, DB 28, Length 720;
Best Local Similarity 97.2%, Pred No. 0.00e+00;
Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGTCCTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 94
|||||
Qy 4547 AGTCCTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4488
|||||
Db 95 CTGCTCTCGGGAAGCGGGCGCTTCTCTAGTCTACGCTGTAGTATCTCTAGTTCGCTG 154
|||||
Qy 4487 CTGCTCTCGGGAAGCGGGCGCTTCTCTAGTCTACGCTGTAGTATCTCTAGTTCGCTG 4428
|||||
Db 155 TAGTGCTGCTGCTCCCAAGTGGGCTGTGTGCACGAACCCCGCTTCAGCCGCGCGCTGC 214
|||||
Qy 4427 TAGTGCTGCTGCTCCCAAGTGGGCTGTGTGCACGAACCCCGCTTCAGCCGCGCGCTGC 4368
|||||
Db 215 GCTTATCCGGTAACCTATGCTGTGAGTCCAAACCCCGGTAAACACGACTTATTCGCCACTG 274
|||||
Qy 4367 GCTTATCCGGTAACCTATGCTGTGAGTCCAAACCCCGGTAAACACGACTTATTCGCCACTG 4308
|||||
Db 275 GCACACACCACTGGTAACAGATTAACAGAGGATATGATAGCGGGTGTCTACAGATTC 334
|||||
Qy 4307 GCACACACCACTGGTAACAGATTAACAGAGGATATGATAGCGGGTGTCTACAGATTC 4248
|||||
Db 335 TTGAAGTGGTGGGCTTAACCTACGGCTACATAGGAACACAGTATTTGGTATCTCGGCTCTG 394
|||||

```

```

Cp 4247 TTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAAGCAGATTTTGGTATCTCGGCTCTG 4188
Db 395 CTGAAGCCAGTTTACCTTCGGGAAAAGAGTTGGTACTCTTGTATCCGCGCAACAAACACC 454
|||||
Cp 4187 CTGAAGCCAGTTTACCTTCGGGAAAAGAGTTGGTACTCTTGTATCCGCGCAACAAACACC 4128
|||||
Db 455 GCTGTAGTGGGTG-TTTTTTTTGGTGAAGCAGCAGATTACGGCGCAGAGAAAAAGGATCT 513
|||||
Cp 4127 GCTGTAGTGGGTGTTTTTTTTTGGTGAAGCAGCAGATTACGGCGCAGAGAAAAAGGATCT 4068
Db 514 CAAGAAGATCTCTTGTATCTTT-CTACGCGGCTCTGAGCGCTAGTGAAGAAAAAATCTCAG 572
|||||
Cp 4067 CAAGAAGATCTCTTGTATCTTTTACGGGGTCTGACGCTAGTGAAGAAAAAATCTCAG 4009
Db 573 TTAAGGGATTTTGGTCATGAGATTATCCAAAGGATCTTCCACTAGATCTTT-AAATTA 631
|||||
Cp 4008 TTAAGGGATTTTGGTCATGAGATTATCCAAAGGATCTTCCACTAGATCTTT-AAATTA 3949
Db 632 AAATGAAGTTTAAATCCACTTAAAGTATATATGATGATGATGATGATGATGATGATGATGAT 690
|||||
Cp 3948 AAATGAAGTTTAAATCCACTTAAAGTATATATGATGATGATGATGATGATGATGATGATGAT 3889
Db 691 TGGCTTATCATGANGGAC-TATCTCAGC 719
|||||
Cp 3888 ATGCTTAATCAGTGAAGCAGCTATCTCAGC 3859

RESULT 11 AG006062 720 bp DNA GSS 14-MAR-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic
DEFINITION survey sequence.
ACCESSION AG006062
NID g2960469
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 762015SpN14.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published only in Database (1998) In press
REFERENCE 2 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1998) to the DBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228 Japan
(E-mail:hattori@ngs.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)

FEATURES
source
1 720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="762015SpN14"
/map="21q"
BASE COUNT 172 a 178 c 170 g 186 t 14 others
ORIGIN
Query Match 12.2%, Score 624, DB 26, Length 720;
Best Local Similarity 97.2%, Pred No. 0.00e+00;
Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGTCCTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 94
|||||
Cp 4547 AGTCCTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4488
|||||
Db 95 CTGCTCTCGGGAAGCGGGCGCTTCTCTAGTCTACGCTGTAGTATCTCTAGTTCGCTG 154
|||||
Cp 4487 CTGCTCTCGGGAAGCGGGCGCTTCTCTAGTCTACGCTGTAGTATCTCTAGTTCGCTG 4428
|||||
Db 155 TAGTGCTGCTGCTCCCAAGTGGGCTGTGTGCACGAACCCCGCTTCAGCCGCGCGCTGC 214
|||||

```



```

*****
Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

*****
MParch_n n.a. - n a database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 19:51:20 1999. MasPar time 8241.22 Seconds
Tabular output not generated. 1546 163 Million cell updates/sec

Title: >US-09-020-716-6
Description: (1-5392) from US09020716.seq
Perfect Score: 5392
N.A. Sequence: 1 CTAAATTTGTAAGCGTTAATA ..... ATTTCCTCGGAAAGTGGCAC 5392
Comp: GATTTAACATTGGCAATTAT ..... TAAAGGGCGCTTTCACGGTG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_to 14:em_un 15:em_vi

Database:
16:gb_bal 17:gb_ba2 18:gb_btg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi

Statistics: Mean 12.732: Variance 8.276: scale 1.538

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```

Result No.	Query Match	Score	Length	ID	Description	Pred No
1	2224	41.3	2961	32	APR2KSP	0.00e+00
2	2228	41.3	2961	32	APR2KSM	0.00e+00
3	2228	41.3	2961	32	APR2KSM	0.00e+00
4	2228	41.3	2961	32	APR2KSM	0.00e+00
5	2228	41.3	2961	32	APR2KSM	0.00e+00
6	2228	41.3	2961	32	APR2KSM	0.00e+00
7	2228	41.3	2961	32	APR2KSM	0.00e+00
8	2228	41.3	2961	32	APR2KSM	0.00e+00
9	2228	41.3	2961	32	APR2KSM	0.00e+00
10	2228	41.3	2961	32	APR2KSM	0.00e+00
11	2228	41.3	2961	32	APR2KSM	0.00e+00
12	2228	41.3	2961	32	APR2KSM	0.00e+00
13	2228	41.3	2961	32	APR2KSM	0.00e+00

C	14	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	15	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	16	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	17	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	18	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	19	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	20	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	21	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	22	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	23	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	24	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	25	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	26	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	27	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	28	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	29	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	30	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	31	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	32	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	33	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	34	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	35	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	36	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	37	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	38	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	39	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	40	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	41	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	42	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00
C	43	2224	41.2	2967	32	U02449	Cloning vector pBR321	0.00e+00
C	44	2224	41.2	2967	32	U02449	Expression vector pBR321	0.00e+00
C	45	2224	41.2	2967	32	U02449	Sequence 26 from Paten	0.00e+00

ALIGNMENTS

RESULT	1	APR2KSP	2961 bp	DNA	circular	SYN	10-MAY-1995
LOCUS		pBluescript II KS(-)		vector DNA			excised from lambda
DEFINITION		ZAPII					
ACCESSION		X53327					
NID		958061					
KEYWORDS		artificial sequence, cloning vector, expression vector, vector, cloning vectors					
SOURCE		cloning vectors					
ORGANISM		Artificial sequences: Cloning vectors.					
REFERENCE		1 (bases 1 to 2961)					
AUTHORS		Thomas, E.A.					
TITLE		Direct Submission					
JOURNAL		Submitted (20-FEB-1990) to the EMBL/GenBank/DBJ databases. Thomas E.A., Stratagene Clonin Systems, 11099 North Torrey Pines Rd., La Jolla, CA 92037, USA					
REFERENCE		2 (bases 1 to 2961)					
AUTHORS		Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.					
TITLE		Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties					
JOURNAL		Nucleic Acids Res. 16 (15), 7583-7600 (1988)					
MEDLINE		88319944					
REFERENCE		3 (bases 1 to 2961)					
AUTHORS		Altting-Wees, M.A. and Short, J.M.					
TITLE		pBluescript II: gene mapping vectors					
JOURNAL		Nucleic Acids Res. 17 (22), 4444 (1989)					
MEDLINE		90067967					
FEATURES		Location/Qualifiers					
source		1..2961					
misc_feature		1..2961					
BASE COUNT		747 a 738 c 755 g 721 t					
ORIGIN							
Query Match		41.3%					
Best Local Similarity		100.0%					
		Score 2228: DB 32: Length 2961:					
		Pred. No. 0.00e+00:					

Matches	2248	Conservative	0	Mismatches	0	Indels	0	Gaps	0
DB	734	GTCAGCTGAGAGGGGGGGGGTATGACGCTTTTCTGCTTAATGAGGGTTAAATGC	793						
QY	4165	GTCAGCTGAGAGGGGGGGGGTATGACGCTTTTCTGCTTAATGAGGGTTAAATGC	3224						
DB	744	GCGTTGGGTAATCATGAGTCATAGCTTTCTGCTGTGAAATGTTTATCGGCTCACAAT	853						
QY	4226	GCGTTGGGTAATCATGAGTCATAGCTTTCTGCTGTGAAATGTTTATCGGCTCACAAT	3384						
DB	864	TATATACATACATAGAGAGAGAAATATAAATCTTAAATCTGAGGGGGCTTAATAGTAGG	913						
QY	286	TATATACATACATAGAGAGAGAGAAATATAAATCTTAAATCTGAGGGGGCTTAATAGTAGG	3344						
DB	914	TAAATCAGATTAATTGGGTTTGGGTCACCTGGGGGTTTCTAACTGGGAAACCGCTGGG	973						
QY	4346	TAAATCAGATTAATTGGGTTTGGGTCACCTGGGGGTTTCTAACTGGGAAACCGCTGGG	3404						
DB	974	CTAGCTGGATTAAATCAATGGTAAATCTGGGGGAGAGGCTTTTGGTATTTGGGGCTC	1033						
QY	4405	CTAGCTGGATTAAATCAATGGTAAATCTGGGGGAGAGGCTTTTGGTATTTGGGGCTC	3464						
DB	1034	TTGGGTTTCTGGTCTCATCTAGCTATGCTGTGCTGCTTTGCTTGGGGGAGGCGTATC	1093						
QY	4465	TTGGGTTTCTGGTCTCATCTAGCTATGCTGTGCTTTGCTTGGGGGAGGCGTATC	3524						
DB	1094	ATCTCAGCTCAAGGGGGTAAATAGCTTATCTCATAGATCTAGAGGATACCGCAGCAAGAA	1153						
QY	4626	ATCTCAGCTCAAGGGGGTAAATAGCTTATCTCATAGATCTAGAGGATACCGCAGCAAGAA	3584						
DB	1154	CATCTGAGCAAAAAGCTAGCAAAAAGCAAGAAAGCTTAAAAACCGCGCTTTGGTGGCGTT	1213						
QY	4686	CATCTGAGCAAAAAGCTAGCAAAAAGCAAGAAAGCTTAAAAACCGCGCTTTGGTGGCGTT	3644						
DB	1214	TTTGTATAGTTTGGTCTTTTCTTTTATATATATATACAAAAATGAGAGCTTAAAGAGAGTGC	1273						
QY	4646	TTTGTATAGTTTGGTCTTTTCTTTTATATATATATACAAAAATGAGAGCTTAAAGAGAGTGC	3704						
DB	1274	GTGAAACCGCAGCAGCTATATAATCATATACAGGCTTTTCTGCTGAGCTTCTGCTGGG	1333						
QY	4706	GTGAAACCGCAGCAGCTATATAATCATATACAGGCTTTTCTGCTGAGCTTCTGCTGGG	3764						
DB	1334	CTTTCGCTTTTGTGAGTTTGTGCTTATAGGATAGCTGTGGGGCTTTTCTGCTTGGGGAAG	1393						
QY	4766	CTTTCGCTTTTGTGAGTTTGTGCTTATAGGATAGCTGTGGGGCTTTTCTGCTTGGGGAAG	3824						
DB	1394	CTTTCGCTTTTGTGAGTTTGTGCTTATAGGATAGCTGTGGGGCTTTTCTGCTTGGGCTC	1453						
QY	4826	CTTTCGCTTTTGTGAGTTTGTGCTTATAGGATAGCTGTGGGGCTTTTCTGCTTGGGCTC	3884						
DB	1454	CAAGCTTGGGCTTGTGGCAGCAAGGCTTTTATAGGGGAGCTTGTGGGCTTTATCGGTA	1513						
QY	3886	CAAGCTTGGGCTTGTGGCAGCAAGGCTTTTATAGGGGAGCTTGTGGGCTTTATCGGTA	3944						
DB	1514	CTATGCTTCTGAGCTTAACTGGGTAAGAAATATATGGGATGGCAGGAGCCACTGG	1573						
QY	4946	CTATGCTTCTGAGCTTAACTGGGTAAGAAATATATGGGATGGCAGGAGCCACTGG	4004						
DB	1574	TAAACGATTACGACAGGATGATATAGGCTGTGTATCAATGCTTTCTGAAGTGGTGGCC	1633						
QY	4006	TAAACGATTACGACAGGATGATATAGGCTGTGTATCAATGCTTTCTGAAGTGGTGGCC	4064						
DB	1634	TAACTAGGCTACACTAGAAAGACATTTTGGTATCTGGGCTCTGCTTGAAGGACAGTTAC	1693						
QY	4066	TAACTAGGCTACACTAGAAAGACATTTTGGTATCTGGGCTCTGCTTGAAGGACAGTTAC	4124						
DB	1694	CTTTCGAAAAAAGCTTTGCTATCTTCTGCACTGGCAAAATGAGTATGCTAGCTGG	1753						
QY	4126	CTTTCGAAAAAAGCTTTGCTATCTTCTGCACTGGCAAAATGAGTATGCTAGCTGG	4184						
DB	1754	TTTTTTTGTGTCAGCAATATATATAGGCGCAGAAAAAGGATGCTCAAGAAGATCGTTT	1813						
QY	4126	TTTTTTTGTGTCAGCAATATATATAGGCGCAGAAAAAGGATGCTCAAGAAGATCGTTT	4244						

[illegible]

RESULT 4 XX035136 4289 bp DNA circular SYN 26-SEP-1995
DEFINITION Plasmid pBSL97 cloning vector, complete sequence.
ACCESSION U35136
KEYWORDS NID
KEYWORDS 994924
SOURCE cloning vector pBSL97.
ORGANISM Cloning vector pBSL97.
REFERENCE 1 (bases 1 to 4289)
AUTHORS Alexeyev M.F., Shokolenko I.N. and Croughan T.P.
TITLE Improved antibiotic resistance gene cassette and amara elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis
JOURNAL Gene 160 (1), 63-67 (1995)
REFERENCE 2 (bases 1 to 4289)
AUTHORS Henden P.N.
TITLE Direct Submission
JOURNAL Submitted (31-AGO 1995) Paul N. Henden, Laboratory of Mathematical Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES Location/Qualifiers
 1..4289
 /organism="Cloning vector pBSL97"
 /plasmid="pBSL97"
 /db_xref="taxon:42709"
 complement(804..1598)
 /codon_start=1
 /transl_table=1
 /product="mex-xylin phosphotransferase"
 /db_xref="pid:994924"
 /transl_table="MIFODGTHADSPAAKVERIFGVYDWAQIIGRSDAVFRLSAGRP
 PVLKVTDLSCALNELODEARLSWLAATGVPCAAVLVDVTEAGRDLLGVEPQDGL
 LSSHLAPAEKVSIMADARRCHTICATGFFGHCARHTEAFTEARLMEALVQDGLDE
 EQHGLAFAPFAPKAPMPDSEGVVTHGPAITPNIMVNGPESFIDCGRUGVADRY
 CGGLATRDCAEEFGKMAAGFFVLVYIAAFISQIFAFYELLEFF"
 complement(3301..4161)
 /EC_number="3.5.2.6"
 /codon_start=1
 /transl_table=1
 /function="ampicillin resistance"
 /product="beta lactamase"
 /db_xref="pid:994925"
 /transl_table="MSICHPFVALIPFAAPFLPFAHPETLVKVKDAEDFQGARVGY
 IELGNSKILLESFPEPFEPFMSTIEKVLIGAVLSKIDAGQGEQPRRIHYSQNDLVE
 YSPVTEKHITGGMTVRELCAATMSNTAANLLITISGKELTAEFLNMHQRVTRI
 DRWPELNEALPNDRDTTPVAMATLRLKLTGDELTILASROOLDIMMEADKVGPL
 LRSLAPAGWFLAUKSGAFERPSRSTIAATSPGKPSPIVVVITISQATMDERNPQIA
 EIGASLIKH"
 BASE COUNT 1024 a 1157 c 1111 g 997 t
 ORIGIN
 Query Match 41.3% Score 2229; DB 32; Length 4289;
 Best Local Similarity 100.0% P-adj No. 0.000000
 Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 2062 GTGAGCTTCAAGCGGCGGCGGATGACGAGCTTTTGGTCTTACGAGCGGTTAATTCG 2121
 QY 4165 GTGAGCTTCAAGCGGCGGCGGATGACGAGCTTTTGGTCTTACGAGCGGTTAATTCG 3224
 Db 2122 GGTCTTGGTCTTAAATATGAGTATATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 2181
 QY 4225 GGTCTTGGTCTTAAATATGAGTATATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 3284
 Db 2182 TGCACACATACACGCGGCGGCGGATGACGAGCTTTTGGTCTTACGAGCGGTTAATTCG 2241
 QY 4285 TGCACATATACAGCGGCGGCGGATGACGAGCTTTTGGTCTTACGAGCGGTTAATTCG 3344
 Db 2242 GTAACTGCAATTAATTTGCTTGGTCTTAAATATGAGTATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 2301
 QY 4345 GTAACTGCAATTAATTTGCTTGGTCTTAAATATGAGTATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 3404

Db 2302 CTACCTGCTATTAATGAATCGGCTAAATCGGCTAAATCGGCTAAATCGGCTATTCGAGCTTC 2361
 QY 3406 CTACCTGCTATTAATGAATCGGCTAAATCGGCTAAATCGGCTAAATCGGCTATTCGAGCTTC 4464
 Db 2362 TTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2421
 QY 3465 TTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4524
 Db 2422 AGCTCATCTAAAGCGGCTAAATCGGCTATTCGAGCTTCGAGCTTCGAGCTTCGAGCTTC 2481
 QY 3525 AGCTCATCTAAAGCGGCTAAATCGGCTATTCGAGCTTCGAGCTTCGAGCTTCGAGCTTC 4584
 Db 2482 CATGTGACAAAGCGGCTAAATCGGCTATTCGAGCTTCGAGCTTCGAGCTTCGAGCTTC 2541
 QY 3585 CATGTGACAAAGCGGCTAAATCGGCTATTCGAGCTTCGAGCTTCGAGCTTCGAGCTTC 4644
 Db 2542 TTTCATAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2601
 QY 3645 TTTCATAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 4704
 Db 2602 GCGAAATCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2661
 QY 3705 GCGAAATCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 4764
 Db 2662 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2721
 QY 3765 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4824
 Db 2722 GCGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2781
 QY 3825 GCGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4884
 Db 2782 CAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2841
 QY 3885 CAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4944
 Db 2842 CTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2901
 QY 3945 CTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4004
 Db 2902 TAACAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2961
 QY 4005 TAACAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4064
 Db 2962 TAATAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4121
 QY 4065 TAATAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4184
 Db 3022 CTTCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3081
 QY 4125 CTTCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4184
 Db 3082 TTTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4141
 QY 4185 TTTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4244
 Db 3142 GATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4201
 QY 4245 GATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4304
 Db 3202 CATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4261
 QY 4305 CATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4364
 Db 3262 ATCAATCTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4321
 QY 4365 ATCAATCTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4424
 Db 3322 GGTCTTGGTCTTAAATATGAGTATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 3381
 QY 4425 GGTCTTGGTCTTAAATATGAGTATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 4484
 Db 3382 GTAGACTCAATTAATTTGCTTGGTCTTAAATATGAGTATATTTCTGCTGAGGAAATGGTATGCGCTACAAAT 4441

QY	4485	GTAGAAGCTACGATACGGAGGGGTTAGATATCTGGGACCAAGTCTGCAATATGATACGGG	4544
Db	3442	ATACCTAACCTCACCAGGCTTCAGATTATATCAGCAATAAACCAGCACGCCGAAGGCCCGA	3501
QY	4545	AGACCCACGCTCACGGGCTTCAGATTATACAGCAATAAACCAGCACGCCGAAGGCCCGA	4604
Db	3502	ACGCCAAGCTGTGTCTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGCCGGGA	3561
QY	4605	GCGCAAGATGTGTGTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGCCGGGA	4664
Db	3562	AGCTCAGTAAGTAGTTCGCCAGTTTAATAGTTTGGCAACGTTCTGCCAATTCGTCACAGG	3621
QY	4665	AGCTCAGTAAGTAGTTCGCCAGTTTAATAGTTTGGCAACGTTCTGCCAATTCGTCACAGG	4724
Db	3622	CATCGTGTGTCCAGCTCGTCTGTTGGTATGGCTTCATCAGCTCCGGTTCGCAACGATC	3681
QY	4725	CATCGTGTGTCCAGCTCGTCTGTTGGTATGGCTTCATCAGCTCCGGTTCGCAACGATC	4784
Db	3682	AAGGCAAGTATATGATCCGCACTGTTGTCGCAAAAACCCGTTACCTCTCTGGTCTCTC	3741
QY	4785	AATGCAAGTATATGATCCGCACTGTTGTCGCAAAAACCCGTTACCTCTCTGGTCTCTC	4844
Db	3742	GATGCTGTGCAGAAATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCA	3801
QY	4845	GATGCTGTGCAGAAATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCA	4904
Db	3802	TAACTCTCTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC	3861
QY	4905	TAACTCTCTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC	4964
Db	3862	CAAGTCAATCTGAGAAATGTTATGCGCGCAGCGAGTTGCTCTTGGCGCGGCTCAATACG	3921
QY	4965	CAAGTCAATCTGAGAAATGTTATGCGCGCAGCGAGTTGCTCTTGGCGCGGCTCAATACG	5024
Db	3922	GTATAATAGCGGCAATAGCAGAACTTTAAAATGTGCTCATCATTTGGAAAAAGTGTCTC	3981
QY	5025	GGATAATAGCGGCAATAGCAGAACTTTAAAATGTGCTCATCATTTGGAAAAAGTGTCTC	5084
Db	3982	GSGCGCAAACTCTCAAGGATCTTACCGGTGTTGAGATCCAGTTTCGATGTAAOCCACTCG	4041
QY	5085	GSGCGCAAACTCTCAAGGATCTTACCGGTGTTGAGATCCAGTTTCGATGTAAOCCACTCG	5144
Db	4042	TGCACCCAACTGATCTTCAGCATCTTTTACITTCACCAAGCGTTTCTGGGTGAGCAAAAC	4101
QY	5145	TGCACCCAACTGATCTTCAGCATCTTTTACITTCACCAAGCGTTTCTGGGTGAGCAAAAC	5204
Db	4102	AGGAAGCAAAATGCGGCAAAAAGGGAATAAGGCGCAGACGGAAATGTTGCAATACTCAT	4161
QY	5205	AGGAAGCAAAATGCGGCAAAAAGGGAATAAGGCGCAGACGGAAATGTTGCAATACTCAT	5264
Db	4162	ACTCTTCTCTTTTCAATATATTGAAGCATTTATCAGCGTTTATGTCTCATGAGCGGATA	4221
QY	5265	ACTCTTCTCTTTTCAATATATTGAAGCATTTATCAGCGTTTATGTCTCATGAGCGGATA	5324
Db	4222	CATATTCAATATATTAGAAAAATAAAATAAATAGGTTTCCGCGCACATTTCGCGGAAA	4281
QY	5325	CATATTGAATGATTATAGAAAAATAAACAATAAGGGTTCGCGCGCACATTTCGCGGAAA	5384
Db	4282	AGTCCAC 4289	
QY	5385	AGTCCAC 5392	

```

artificial sequence.
1 (bases 1 to 4412)
  AUTHORS      Liu,L., Dasgupta,I., Davies,J.W. and Hull,R.
  TITLE        Modified vectors for monocot transformation toward virus resistance
  JOURNAL      Unpublished
2 (bases 1 to 4412)
  McElroy,D., Blowers,A.D., Jones,B. and Wu,R.
  Construction of expression vectors based on the rice actin 1 (Act1)
  5' region for use in monocot transformation
  Mol. Gen. Genet. 231 (1), 150-160 (1991)
92092956
(sites)
3 (bases 27 to 282)
  Depicker,A., Stachel,S., Phageol,S., Zartzyev,I.P. and Goodman,H.M.
  Nopaline synthase: transcript mapping and DNA sequence
  J. Mol. Appl. Genet. 1 (6), 561-573 (1982)
83110651
4 (bases 1 to 4412)
  Liu,L.
  Direct Submission
  Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus
  Research, Colney Lane, Norwich, United Kingdom, NR4 7UH
  Location/Qualifiers
    1..4412
      /organism="expression vectors"
      /db_xref="taxon:33776"
    1..30
      /note="Multiple cloning sites"
    3..5
      /note="start codon for gene expression"
    27..282
      /note="nos terminator"
      /citation={3}
    506..962
      /note="phage fl region"
    1093..11953
      /codon_start=1
      /transl_table=11
      /product="beta-lactamase"
      /db_xref="PID:g452345"
      /translation="MSIQHPVALIDFFAAFLPVAHPETLVKVKAEQGLGAEVY
      YELRNSKILESEFPEERPMWTFKVIQAVLSPIQAGDELLGPIRISQALVEL
      VSPYRHLRIDIVRELCGAAITMSDNTAANLLITGGPKETAFHNGDRHVTLE
      DWPELENAIPNEDRITTPVAMATLTSLITGELLTSLSPQOLIDWVFADKVGAGP
      LRSALPAGWFTADKSGAGPGSGCIIAALGPDGKPSPIWVYITGSGQATMDERNQRIA
      EIGASLIKHW"
    1954..2894
      /note="ColEI origin"
    3181..4412
      /note="rice actin 1 5' region (act1 promoter)"
      /citation={2}
BASE COUNT 1127 a 1107 c 1036 g 1142 t
ORIGIN

Query Match      41.3%   Score 2225; DB 32; Length 4412;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2225, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Db 965 GTGGCACTTTTGGGGAATGTGGCGGAACCCCACTATTGTTTATTITTTCTAATAACATT 1024
Cp 5392 GTGGCACTTTTGGGGAATGTGGCGGAACCCCACTATTGTTTATTITTTCTAATAACATT 5333
Db 1025 CAATATGATCGGTCATGAGACAAATACCCCTGATGAATGCTTCAAATATGAAAAA 1084
Cp 5332 CAATATGATCGGTCATGAGACAAATACCCCTGATGAATGCTTCAAATATGAAAAA 5273
Db 1085 GGAAGAGTATGAGTATTCAAACATTTCCCTGTGGCGCCCTATTCCCTTTTTTGGCGCAATTT 1144
Cp 5272 GGAAGAGTATGAGTATTCAAACATTTCCCTGTGGCGCCCTATTCCCTTTTTTGGCGCAATTT 5213
Db 1145 GCCTTCCTGTTTTTGGTCACCCAGAAACGCTGTGTGTAAGTAAGAATGCTGGAAGATCAGT 1204
Cp 5212 GCCTTCCTGTTTTTGGTCACCCAGAAACGCTGTGTGTAAGTAAGAATGCTGGAAGATCAGT 5153

```



```

artificial sequence; cloning vectors.
1 (bases 1 to 9655)
AUTHORS Mazzarella, R. and Pillia, G.
TITLE Recombination trapping: an 'in vivo' approach to recover cDNAs
encoded in YACS
JOURNAL Unpublished (1997)
COMMENT GSDB:S:1274444.
FEATURES
    Location/Qualifiers
        source
            1..9655
                /organism="Cloning vector pGR8"
                /plasmid="Cloning vector pGR8"
                /db_xref="taxon:64712"
        misc_feature
            1..135
                /standard_name="telomere"
        prim_transcript
            145..2097
                /gene="HIS5"
        gene
            145..2097
                /gene="HIS5"
        CDS
            761..1915
                /gene="HIS5"
                /EC_number="2.6.1.9"
                /codon_start=1
                /transl_table=11
                /product="histidinal phosphate aminotransferase"
                /db_xref="PID:g2338009"
                /translation="MVFDLRIVRPKIYNLEPYRCARDFTTEGILLDANENAHGPTPV
                ELKNTLHRYDPDPHOLEFKTAMTKYPNKTSYVANDPEVKPLTADNLCGLVGSDESDA
                IIPACCVREKILVLPPYSSMSSVCANINDIEVCCPLIVSDGSFQMDTEAVLTILKN
                DSIKLKMFVTSNGNTGAKIKTSLIEKVLQWNGLVVDEAYVDFCGGSTAPLVTKY
                PNLVLTQUSKSGFAGIRLGMIAECLRLNMAKAPYNISLASLEYALKAVQDSN
                LKMEATSIKINEEKLKLLKELTALDYDDQVGGDLNELLIRINGDNVLAKKLYV
                QKATQGVVVRFRGELGCGCLRTIVGTGTHEENTHLIKYFKETLYKLANE"
        prim_transcript
            2138..6708
                /gene="LYS2"
        gene
            2138..6708
                /gene="LYS2"
                /EC_number="1.2.1.31"
                /codon_start=1
                /transl_table=11
                /product="alpha-aminoadipate reductase"
                /db_xref="PID:g2358310"
                /translation="MTNEKWIIEKLDNPTLSLPHDFLRPOQEPYTKQATYSVLQPLP
                DVPHDSFNKVAVALSVMAALIYRVTGDDIVLYIANKNLIKRFNIQPTWFNELYSTI
                NNELKLSIEANFSDELAEKIQSCDLERTPOLFRFLAFLNODFKLDEFKHLVDF
                ALNLDTSNAHVLNLIYNSLLYSNERVTIVADQFTQYLTAALSDPNCILIKSLITAS
                SKDSLDPKTLNGLWDFVGCIHDFIQDNAEAPPERTCVETPTLNSKRSKSFITRDIN
                RTSNVDHPTIKLIGIKRGEVVMYIYSSRGVDMVCMVGLKAGATFSVIDPAPPAROT
                IYLVAKRPLGLIYRAAGQDLQVEDYINDELEIVSRINSIAIQENGTIEGKLNDGE
                DVLPADHYKDTRTGVVGGPSNPTLSFTSGSEGIKPGVLGRHFSLAYEFNWSKREN
                LTENDKFTLSGIAHDPIQDMFTPLFLGAOLYVPTQDDIGTPGRSLAEMWSKYCTVT
                HLTPMAGQLTLQAATPTPKLHAFFVGDILTFRDCLRLQTLAENCRIVNMYGTETQ
                RAVSEYFEVAKNDNDNFLLKKLDVMPACKGMLNVLQVLLVNRNDRTOICGIEIGIYV
                RAGLAEGYKGLPELNKEKFEVNNVEVDKHNVLNDNGEPWQFMRFLGRDLYRTGD
                LGRYPNGCECCGRADDOQVKIRGFEIIEUGEDITHISOHLVPEINLTLRKNDNED
                LITFMYPRDKLQSKFQSDPKQEVETDPIVKGLIYHLLSKDIRFLPKRLASVAM
                PSIIYVMDKLPLNPGKVDKPLQETPKQLNVAENTVSETDSDQFTNVREVRDLW
                LSLPTPKASVSPDQSFVFDLGSILATKMIPTLKKLQVPLPLGTIFPKYPTIKAFAA
                EIDRIKSSGSSQGEVENVTANYAEDAKKLVETLPSYPSPREYFVSPNSAGKTTIN
                VFYGTGTGLSYILADLLGRSPKNYSFKFAHVRADKEEAFARGLKQYITGWNE
                KFAINSVLVGLDQSKSGFGLSDEKMDLANTVDIIHNGALVHWYYPYAKLRDPNVIS
                TINMSLAAGKPKDFEVSSTILDTEYFNLSKLVSEKGPGLILESDLLMSASGL
                TGGYGSQWAAEYIIRRAGERGLGCIVRPGYVTVGASANGSNTDDFLRLKGSVL
                GKIPDTENSVMVDPVAVRVVATSLNPPKENELAVAVTGHPRILFKDYLYTLHDY
                GYDVEIESYKMKKSLSEASVIDRNEENALYPLLHMVDNLPESTKAPELDRNVAVL
                KKDANTGVDSWNGICVTPPEEVGIVIAFLNKVGLFPPTTHNDKPLPSIELTQAISL
                VASGAGAGSSAAA"
        misc_feature
            complement(6709..7185)
        gene
            7186..7496
                /gene="Lac2"
        prim_transcript
            7186..7496

```

```

/misc_feature
    /gene="Lac2"
    7339..7445
    /gene="Lac2"
    /standard_name="multiple cloning sites"
    complement(7910..8030)
    /standard_name="plasmid ORI"
    complement(8659..9522)
    /gene="AMP"
    complement(8659..9519)
    /gene="AMP"
    /function="ampicillin resistance"
    /codon_start=1
    /transl_table=11
    /product="beta-lactamase"
    /db_xref="PID:g2358011"
    /translation="MSIQHFRVALIPEFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
    IELNLSGKILSEPRPEEFPMSTFRKVIQGVAVLSDITAEQGLDPRRIHYSNLSLVE
    YSPVTEKHTDGMTVRELCSAAITMSNTAANLLITIGGKELTALFLHMGDHTVRL
    DRWPELNEAIPNDERDTIMPVMAITLRLKLLTGELLILASRQQLIDMDEADKVAQPL
    LPSALPAGWFIADKSGAGPGSPGIIAALGPDGKPSPIVVIYITGSGATWCEENPQIA
    EIGASLIKHW"
    prim_transcript complement(8734..9522)
    /gene="AMP"
BASE COUNT 2804 a 2009 c 2099 g 2743 t
ORIGIN
Query Match 41.3% Score 2228; DB 32; Length 9655;
Best Local Similarity 100.0%; Pred. NO. 0.00e+00;
Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 7420 GTGAGCTCGAGGGGGGGGATCCAGAGTITTTTTCCTTTTAAAGAGGTTAAATTCG 7479
Qy 3165 GTGAGCTCGAGGGGGGGGATCCAGAGTITTTTTCCTTTTAAAGAGGTTAAATTCG 3224
Db 7480 GCCTTGGGGTAAATCAATGATAGTCTGTTTCTCTGTGTAATTTTATTCGCTCAAT 7539
Qy 3225 GCCTTGGGGTAAATCAATGATAGTCTGTTTCTCTGTGTAATTTTATTCGCTCAAT 3284
Db 7540 TCCACACACATACGAGCGGGAAGGATAAAGTGTAAAGCTTGGGGTGGGAGAGAG 7599
Qy 3285 TCCACACACATACGAGCGGGAAGGATAAAGTGTAAAGCTTGGGGTGGGAGAGAG 3344
Db 7600 CTAATTCATTAATTTTTCCTTCATCTGCTGCTGCTTTCACAGTGGGAAACCTGG 7659
Qy 3345 CTAATTCATTAATTTTTCCTTCATCTGCTGCTGCTTTCACAGTGGGAAACCTGG 3404
Db 7660 CCAGCTGCATTAAATGAATCGGCAACGGCGGGGAGAGAGCGGTTTGGCTATTCGG 7719
Qy 3405 CCAGCTGCATTAAATGAATCGGCAACGGCGGGGAGAGAGCGGTTTGGCTATTCGG 3464
Db 7720 TTCCGCTTTCCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7779
Qy 3465 TTCCGCTTTCCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3524
Db 7780 AGTCTACTCAAGGGCGGTAAATCGGTTATTCACACAAATCAGGAGGATAACAGCAG 7839
Qy 3525 AGTCTACTCAAGGGCGGTAAATCGGTTATTCACACAAATCAGGAGGATAACAGCAG 3584
Db 7840 CATGTGAGCAAAAGGCCACCAAAAGCCAGAACCGTAAAAAGCGCGGCTTCTGCTGCT 7899
Qy 3585 CATGTGAGCAAAAGGCCACCAAAAGCCAGAACCGTAAAAAGCGCGGCTTCTGCTGCT 3644
Db 7900 TTCCCATAGGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7959
Qy 3645 TTCCCATAGGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3704
Db 7960 GCGAAACCCGACAGGACTATAAGATACACAGCGGTTTTCGCTTGGAGCTTCTGCTGCT 8019
Qy 3705 GCGAAACCCGACAGGACTATAAGATACACAGCGGTTTTCGCTTGGAGCTTCTGCTGCT 3764
Db 8020 CTCTCTGTTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTCTCTCTTCCGAG 8079
Qy 3765 CTCTCTGTTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTCTCTCTTCCGAG 3824

```


Dh 1393 GGAAGATATGATTAACATTTCCGTTGTCGCCCTTATTCCTTTTTCGGCATTTT 1452
Cp 5272 GGAAGATATGATTAACATTTCCGTTGTCGCCCTTATTCCTTTTTCGGCATTTT 5213
Dh 1453 GCCTTCCTGTTTTCCTACCCAGAAACGCTGTTGAAGTAAAGATGCTGAAGTACGT 1512
Cp 5212 GCCTTCCTGTTTTCCTACCCAGAAACGCTGTTGAAGTAAAGATGCTGAAGTACGT 5153
Dh 1513 TGGTGTACAGTGGTGTATCATCGAACTGGATCTCAACAGCGGTAAAGTTCCTTGAGGTT 1572
Cp 5152 TGGTGTACAGTGGTGTATCATCGAACTGGATCTCAACAGCGGTAAAGTTCCTTGAGGTT 5093
Dh 1573 TTCCGCCGGAAGACGTTTTCCTAATGATGAGCATTATTAAGTTCCTGATGTCGCGGG 1632
Cp 5092 TTCCGCCGGAAGACGTTTTCCTAATGATGAGCATTATTAAGTTCCTGATGTCGCGGG 5033
Dh 1633 TATTATCCCTATTGAGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCCTAGA 1692
Cp 5032 TATTATCCCTATTGAGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCCTAGA 4973
Dh 1693 ATGACTTGGTGTAGTATCACTACAGTACAGTAAAGCATCTTACGGATGGCATGACAGTAA 1752
Cp 4972 ATGACTTGGTGTAGTATCACTACAGTACAGTAAAGCATCTTACGGATGGCATGACAGTAA 4913
Dh 1753 GAGAAATATGAGTGTGTCATACCATGATGATTAACACTGCGGCGCAACTTACTTCTGA 1812
Cp 4912 GAGAAATATGAGTGTGTCATACCATGATGATTAACACTGCGGCGCAACTTACTTCTGA 4853
Dh 1813 CAACGATCGGAGGACCAAGAGCACTAACGCTTTTTCACACATGGGGATCATGTAA 1872
Cp 4852 CAACGATCGGAGGACCAAGAGCACTAACGCTTTTTCACACATGGGGATCATGTAA 4793
Dh 1873 CTGCGCTTGATGTTGGAAACCGGAGTGAATGAAGCCATACCAACAGCAGCGTGACA 1932
Cp 4792 CTGCGCTTGATGTTGGAAACCGGAGTGAATGAAGCCATACCAACAGCAGCGTGACA 4733
Dh 1933 CCAGATGCGCTGTAGTAATGGCAACACGTTGGCGCAACTATTAAGTGGCAACTACTTA 1992
Cp 4732 CCAGATGCGCTGTAGTAATGGCAACACGTTGGCGCAACTATTAAGTGGCAACTACTTA 4673
Dh 1993 CTCTAGCTTCCGCGCAACAATTAATAGACTGGATGAGCGGATGAAGTTGCAGGACAC 2052
Cp 4672 CTCTAGCTTCCGCGCAACAATTAATAGACTGGATGAGCGGATGAAGTTGCAGGACAC 4613
Dh 2053 TTCTGCGCTGGCGCTTCCGCGTGGTGTATTAAGTGGATGAGCGGTGAGC 2112
Cp 4612 TTCTGCGCTGGCGCTTCCGCGTGGTGTATTAAGTGGATGAGCGGTGAGC 4553
Dh 2113 GTGGGTCTGCGGTATCATTTGCACTGCGGCGGATGATGATGCGCTCCGCTATCGTAG 2172
Cp 4552 GTGGGTCTGCGGTATCATTTGCACTGCGGCGGATGATGATGCGCTCCGCTATCGTAG 4493
Dh 2173 TTATCTACAGCAGGAGTATGAGTAAAGCAATAGACAGATCGCTGAGA 2232
Cp 4492 TTATCTACAGCAGGAGTATGAGTAAAGCAATAGACAGATCGCTGAGA 4433
Dh 2233 TAGTGCTCTACTGATTAAGCATTTGATGTCAGACCAAGTTTACTCATATATCTTT 2292
Cp 4432 TAGTGCTCTACTGATTAAGCATTTGATGTCAGACCAAGTTTACTCATATATCTTT 4373
Dh 2293 AGATTGATTTAAACTTCACTTTTAAAGGATCTAGGTGAAGATCCCTTTTGATA 2352
Cp 4372 AGATTGATTTAAACTTCACTTTTAAAGGATCTAGGTGAAGATCCCTTTTGATA 4313
Dh 2353 ATCTCATGACCAAAATCCCTTAAGTGTGTTTTCGTTCCACTGAGGCTGAGACCCCGTAG 2412
Cp 4312 ATCTCATGACCAAAATCCCTTAAGTGTGTTTTCGTTCCACTGAGGCTGAGACCCCGTAG 4253
Dh 2413 AAAAGATCAAGGATCTTCTGAGATCCCTTTTTCGCGCGTAATCTGCTGCTTGCAAA 2472
Cp 4252 AAAAGATCAAGGATCTTCTGAGATCCCTTTTTCGCGCGTAATCTGCTGCTTGCAAA 4193
Dh 2473 CAAAAAACCACCTTACAGCGTGGTGTGTTTTCGCGGATCAAGAGTACCAACTCTTT 2532

Cp 4192 CAAAAAACCACCTTACAGCGTGGTGTGTTTTCGCGGATCAAGAGTACCAACTCTTT 4133
Dh 2533 TTCGGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTACTGTAGC 2592
Cp 4132 TTCGGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTACTGTAGC 4073
Dh 2593 CGTAGTTAGGCGCACACITCAAGAACICGTAGACAGCGGTATACATACCTGCTCTGTAA 2652
Cp 4072 CGTAGTTAGGCGCACACITCAAGAACICGTAGACAGCGGTATACATACCTGCTCTGTAA 4013
Dh 2653 TCCTGTTACAGTGGCTGCTGCACTGAGTAAATGCTGCTTACCGGTTGACACAA 2712
Cp 4012 TCCTGTTACAGTGGCTGCTGCACTGAGTAAATGCTGCTTACCGGTTGACACAA 3453
Dh 2713 GAGCATAGTACCGGTAAGCGCGCAGCGTGGGTGACCGGGGCTTCGTGCAACAGC 2772
Cp 3952 GAGCATAGTACCGGTAAGCGCGCAGCGTGGGTGACCGGGGCTTCGTGCAACAGC 3893
Dh 2773 CCAGCTTGGAGGGAAGCACTACAGCACTGAGATACCTACAGCGTGAATGASAAA 2832
Cp 3892 CCAGCTTGGAGGGAAGCACTACAGCACTGAGATACCTACAGCGTGAATGASAAA 3833
Dh 2833 GCGCCAGCTTCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2892
Cp 3832 GCGCCAGCTTCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3773
Dh 2893 CAGCAGTGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2952
Cp 3772 CAGCAGTGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3713
Dh 2953 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTGATGCTGCTGAGGAGGAGGAGG 3012
Cp 3712 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTGATGCTGCTGAGGAGGAGGAGG 3653
Dh 3013 TATGGAAGAACGCGCAGCGGCGCTTTTACGCTTCTGCGCTTTTGTGCGCTTTTG 3072
Cp 3652 TATGGAAGAACGCGCAGCGGCGCTTTTACGCTTCTGCGCTTTTGTGCGCTTTTG 3593
Dh 3073 CTCACATGTTCTTCTGCGCTTATCCCTGATCTGCTGATACCGTATTAACGCTTTG 3132
Cp 3592 CTCACATGTTCTTCTGCGCTTATCCCTGATCTGCTGATACCGTATTAACGCTTTG 3533
Dh 3133 AGTGAGCTGATACCGCTCGCGCAGCGGAGCGAGCGAGCGAGCGAGCGAGCGAGG 3192
Cp 3532 AGTGAGCTGATACCGCTCGCGCAGCGGAGCGAGCGAGCGAGCGAGCGAGG 3473
Dh 3193 AAGCGGAAGCGCCCAATACGCAACCGCTCTCCCGCGCGCTTCGCGGATTCATTAAT 3252
Cp 3472 AAGCGGAAGCGCCCAATACGCAACCGCTCTCCCGCGCGCTTCGCGGATTCATTAAT 3413
Dh 3253 GCAGCTGGCAGCAGCGTTTCCGCACTGGAAAGCGGAGTGAAGCGCAAGTAAATATG 3312
Cp 3412 GCAGCTGGCAGCAGCGTTTCCGCACTGGAAAGCGGAGTGAAGCGCAAGTAAATG 3353
Dh 3313 TGAGTGTAGCTACTCATTAAGCAGCGCGAGGCTTTTACACTTTATGCTTCGCGCTGATGT 3372
Cp 3352 TGAGTGTAGCTACTCATTAAGCAGCGCGAGGCTTTTACACTTTATGCTTCGCGCTGATGT 3293
Dh 3373 TGCTGGAATTTGAGCGGATACAAATTTTACAGAGGAACAGCTATGACCATGATAGG 3432
Cp 3292 TGCTGGAATTTGAGCGGATACAAATTTTACAGAGGAACAGCTATGACCATGATAGG 3233
Dh 3433 CCAAGCGCGCAATTAACCTTCACTAAAGGGAACAAAGTGGGTGAGGAGGAGGAGGAGG 3492
Cp 3232 CCAAGCGCGCAATTAACCTTCACTAAAGGGAACAAAGTGGGTGAGGAGGAGGAGGAGG 3173
Dh 3493 AGGTGAGC 3500
Cp 3172 AGGTGAGC 3165

RESULT 8

Cp 4132 TCCGAGGTAACAGCTGCTTCAACAGAGGAGGAGATACCAAAACATCTTCTAGTACG 4073
 Db 2593 GGTAGTAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2652
 Cp 4072 GGTAGTAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 4013
 Db 2653 TGTGTGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2712
 Cp 4012 TGTGTGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3953
 Db 2713 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2772
 Cp 3952 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3893
 Db 2773 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2832
 Cp 3892 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3833
 Db 2833 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2892
 Cp 3832 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3773
 Db 2893 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2952
 Cp 3772 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3713
 Db 2953 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3012
 Cp 3712 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3653
 Db 3013 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3072
 Cp 3652 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3593
 Db 3073 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3132
 Cp 3592 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3533
 Db 3133 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3192
 Cp 3532 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3473
 Db 3193 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3252
 Cp 3472 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3413
 Db 3253 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3312
 Cp 3412 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3353
 Db 3313 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3372
 Cp 3352 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3293
 Db 3373 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3432
 Cp 3292 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3233
 Db 3433 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3492
 Cp 3232 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3173
 Db 3493 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3500
 Cp 3172 GAGGATAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 3165

RESULT 9
 LOCUS EVU67875 10765 bp DNA circular SYN 16-DEC-1997
 DEFINITION pESP-I yeast expression vector, complete sequence.
 ACCESSION U67875
 MID g2689262
 KEYWORDS

SOURCE ORGANISM
 Expression vector pESP-I
 Expression vector pESP-I
 artificial sequence, expression vectors.
 REFERENCE 1 (bases 1 to 10765)
 AUTHORS Lu.Q., Bauer,J.C. and Greener,A.
 TITLE Using Schizosaccharomyces pombe as a host for expression and purification of eukaryotic proteins
 JOURNAL Gene 200 (1-2), 135-144 (1997)
 MEDLINE 98038984
 REFERENCE 2 (bases 1 to 10765)
 AUTHORS Lu.Q.
 TITLE Direct Submission
 JOURNAL Submitted (12-AUG-1996) R&D, Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA
 REFERENCE 3 (bases 1 to 10765)
 AUTHORS Grafsky,A.
 TITLE Direct Submission
 JOURNAL Submitted (16-DEC-1997) R&D, Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA
 REMARK sequence update by submitter
 COMMENT On Dec 16, 1997 this sequence version replaced 111527192.
 pESP-I yeast expression vector is derived from Bluescript from Stratagene for expression and purification of heterologous genes in Schizosaccharomyces pombe.
 FEATURES
 Location/Qualifiers
 1..10765
 /organism="Expression vector pESP-I"
 /db_xref="taxon:51339"
 1401..2261
 /codon_start=1
 /product="beta-lactamase"
 /db_xref="PID:g1527193"
 /translation="MSIQHFRVALIPFFFAFCLPFAHPRLVFAHPDILVKVDAETGLGARVGY IELDNGSKILESEFPPEPPPMSTKFKVILCGAVLSPIDAGQPLGPPHYISNRLVE YSPYTKHLTDGWTVELQSAITMSDNTAANLLTIGGPKELTAFLHNMGSCHVTR LDRWPELNEAIPNDEROTTPVAMATILKLLTGELLILASROOLDIWEAEAKVAPL LPSAPGAFIAUKSGASGERSPSITAAIGPGKSPFVIVYITIGSEATMGERNGLA EIGASLIKHW"
 5464..6558
 /codon_start=1
 /product="3-isopropylmalate dehydrogenase"
 /db_xref="PID:g2689263"
 /translation="MSAPKRIIVYVPGHVGQETIAEAIKVLKAIISVRNRYKTIENH LGGAADATGVPIDLEALKSKVDAILGAVGSGPKWGTGSPVPPGILKIAKFLQI YANRPFNFASDLSLESPKFAKGTFFVVEIVGVYIFKFFEDTCEYVAMTSE QYTPVEQPTTMAAFMALGHEPPLPIWISDKANVLASSIPWPKTVEETIKAFPEPTIK VUHLIDSAAMILVKNPHTLNGIITISNMFSDIISDEASVIPSGLKLLSASLASLPD KNTAFGLYERCHGSAPDLPKNKNPITAILLSAAMMLKLSLNLPEEGKAIEDAVKKVLD AGIRTGDLGSGNSTIEVGDAVEVKLLA"
 8168..8884
 /codon_start=1
 /product="glutathione S-transferase"
 /db_xref="PID:g1527195"
 /translation="MSPILGYMKIKGLVOPTRIIIEYIPEKVEHLYEPDFGKWNK KFEGLPEPNLPYIIGDVKLTGSMATIFIACKHNMKGCGPKPEKREIEMLEAVLDI RYGVSRVAYSKDPETILKFLSKLPEMLKMFEDRLGKHYLNGDHTHPDMFLYDALD VLYMDPMLCAAFKLVCFKFFPEIAIPQIKLYKSEKVIAMVIGWQATFGGCTGDEPY SVLVPRGSDYKDDDKGSGP"
 BASE COUNT 3276 a 2125 c 2201 g 3163 t
 ORIGIN

Query Match 41.3%; Score 2228; DB 32; Length 10765;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1273 GTGGGACTTTTCGGGAAATGTCGGCGAACCCCTATTGTTTATTTTCTAATAATATT 1332
 Cp 5392 GTGGGACTTTTCGGGAAATGTCGGCGAACCCCTATTGTTTATTTTCTAATAATATT 5333
 Db 1333 CAAATATGATCGGTTCATGAGACAAATACCCCTGATAAATGCTCAATAATATTAATAA 1392
 Cp 5332 CAAATATGATCGGTTCATGAGACAAATACCCCTGATAAATGCTCAATAATATTAATAA 5273

QY 4005 TACAGGATTAGCAGCGAGGTAIGAGGCGGIGCTACAGAGTTCTTGAAGTGGTGCC 4064
Db 1634 TAACAGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTCAAGCCAGTTAC 1693
QY 4065 TAACAGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTCAAGCCAGTTAC 4124
Db 1694 CTTCGGAAGAGTGGTAGCTCTTGATCGGCAACAAACACCGCTGGTAGGGTGG 1753
QY 4125 CTTCGGAAGAGTGGTAGCTCTTGATCGGCAACAAACACCGCTGGTAGGGTGG 4184
Db 1754 TTTTITGTTTGAAGCAGCAGATTAGCGCGCAAAAAAAGAGATCTCAAGAGATCCCTTT 1813
QY 4185 TTTTITGTTTGAAGCAGCAGATTAGCGCGCAAAAAAAGAGATCTCAAGAGATCCCTTT 4244
Db 1814 GATCTTTTACGCGGGTCTGACGCTCAGTGGAAAGCAAACTCACGTTAAGGATTTGGT 1873
QY 4245 GATCTTTTACGCGGGTCTGACGCTCAGTGGAAAGCAAACTCACGTTAAGGATTTGGT 4304
Db 1874 CATGAGATATCAAAAGAGATCTTACCTAGATACCTTTTAAATTAATAAGATTTTAA 1933
QY 4305 CATGAGATATCAAAAGAGATCTTACCTAGATACCTTTTAAATTAATAAGATTTTAA 4364
Db 1934 ATCAATCTAAAGTATATAGTAAAGTCTGCTGACAGTTACCAATGCTTAATCAGTCA 1993
QY 4365 ATCAATCTAAAGTATATAGTAAAGTCTGCTGACAGTTACCAATGCTTAATCAGTCA 4424
Db 1994 GGCACCTATCTACGAGTCTGCTTATTTGCTTATCATCATAGTTGCTGACTCCCGTCTG 2053
QY 4425 GGCACCTATCTACGAGTCTGCTTATTTGCTTATCATCATAGTTGCTGACTCCCGTCTG 4484
Db 2054 GTATATACCTAGTATAGGAGGCTTACCATCTGCGCCAGTGGTGAATGATCCGGG 2113
QY 4485 GTATATACCTAGTATAGGAGGCTTACCATCTGCGCCAGTGGTGAATGATCCGGG 4544
Db 2114 AGATCCAGTCTACGCTCTGAGATTTATCAGCAATAAACACGCGCGGAGGCGCA 2173
QY 4545 AGATCCAGTCTACGCTCTGAGATTTATCAGCAATAAACACGCGCGGAGGCGCA 4604
Db 2174 GCGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGCGCGGA 2233
QY 4605 GCGCAGAGTGGTCTGCAACTTTATCGGCTCCATCCAGTCTATTAAATTTGCGCGGA 4664
Db 2234 AGTTAGATTAAGTATGTCGAGTAAATAGTTTSGCAGCTGTTGCCATGCTTACAGG 2293
QY 4665 AGTTAGATTAAGTATGTCGAGTAAATAGTTTSGCAGCTGTTGCCATGCTTACAGG 4724
Db 2294 CATCGTGTGTACGCTCGCTGTTGGTATGTTTCAATCAGTCCGTTCCCAAGGATC 2353
QY 4725 CATCGTGTGTACGCTCGCTGTTGGTATGTTTCAATCAGTCCGTTCCCAAGGATC 4784
Db 2354 AAGCGAGTTACATGATCCGCTATGTTGTAAGGAGGTTAGCTCTCGGCTCTCC 2413
QY 4785 AAGCGAGTTACATGATCCGCTATGTTGTAAGGAGGTTAGCTCTCGGCTCTCC 4844
Db 2414 CATGTTGTGAGAGTAAAGTTGCGGCTGTTATCAGTCTGTTATGCGGAGCTGCA 2473
QY 4845 CATGTTGTGAGAGTAAAGTTGCGGCTGTTATCAGTCTGTTATGCGGAGCTGCA 4904
Db 2474 TAATTTCTTACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2533
QY 4905 TAATTTCTTACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 4964
Db 2534 CAATCATCTGAGATAGTGTATGCGGCGAGCGAGTTGCTCTTGGCCGCGCTCAATAG 2593
QY 4965 CAATCATCTGAGATAGTGTATGCGGCGAGCGAGTTGCTCTTGGCCGCGCTCAATAG 5024
Db 2594 GGATAATACCGCGCCATAGCAGACTTTTAAAGTGTCTCATCATTTGGAAACGTTCTTC 2653
QY 5025 GGATAATACCGCGCCATAGCAGACTTTTAAAGTGTCTCATCATTTGGAAACGTTCTTC 5084
Db 2654 GGGCGGAAATCTCAAGAGATCTTAAGGCTTGGAGATCCAGTCCAGTGAATCCCACTCG 2713
QY 5085 GGGCGGAAATCTCAAGAGATCTTAAGGCTTGGAGATCCAGTCCAGTGAATCCCACTCG 5144

Db 2714 TGACCCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGTGAGCAAAAC 2773
QY 5145 TGACCCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGTGAGCAAAAC 5204
Db 2774 AGAAGCAAAATGCGCAAAAAGGAAATAAGGCGCACACGAAATGTTGAATACTCAT 2833
QY 5205 AGAAGCAAAATGCGCAAAAAGGAAATAAGGCGCACACGAAATGTTGAATACTCAT 5264
Db 2834 ACTCTTCCCTTTTCAATATATTGAAGCAATTTATCAGGGTATTGCTCTCATGAGCGGATA 2893
QY 5265 ACTCTTCCCTTTTCAATATATTGAAGCAATTTATCAGGGTATTGCTCTCATGAGCGGATA 5324
Db 2894 CATATTGAATGATTAGAAAAATAAACAAATAGGGTTCGCGCACATTTCCCGCAAA 2953
QY 5325 CATATTGAATGATTAGAAAAATAAACAAATAGGGTTCGCGCACATTTCCCGCAAA 5384
Db 2954 AGTGC 2958
QY 5385 AGTGC 5389

RESULT 12

LOCUS SYNBLKSPV 2964 bp DNA circular SYN 04-JUN-1993
DEFINITION Bluescribe KS Plus cloning vector.

ACCESSION L08785

NID 9310729

KEYWORDS

SOURCE Synthetic construct DNA.

ORGANISM

REFERENCE 1 (bases 1 to 2964)

AUTHORS Gilbert W.

JOURNAL

TITLE Obtained from VecBase 3.0

COMMENT

These data and their annotation were supplied to GenBank by Will

Gilbert under the auspices of the GenBank Curator Program.

Bluescribe KS Plus - Cloning vector

ENTRY BLUESKP #TYPE DNA CIRCULAR

TITLE Bluescribe KS Plus - Cloning vector

DATE 28-JAN-1987

#sequence 02-FEB-1987

#sequence 04-MAR-1987

#sequence 03-APR-1987

ACCESSION VB0078

SOURCE artificial

REFERENCE

#number 1

#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Scragg J.

#journal Gene (1987) in press

#citation Sequence data from StrataGene

#comment sequence correction according to StrataGene COMMENT:

Obtained from StrataGene on floppy disc.

1409/10 'Ar' to 'TA' to match revised sequence of pBR322

Revised 4-MAR-1987 to match sequence of pUC19 on request

of StrataGene

Polylinker region revised 03-APR-1987 according to StrataGene

COMMENT

The stand shown corresponds to pUC19c.

As in the published sequence of pUC19c, The M13mp19 lacZ region

is on the complementary strand.

COMMENT

This vector contains the fl origin so that the plus strand

can be obtained upon fl superinfection.

KEYWORDS

CROSSREFERENCE

#parent

VecBase(3):BlueM13p

#parent

VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,

VecSource(3):bGalKS, GenBank(50):PF1


```

QY 4665 AGTACAGTAAGTATTTCGCGAGTTAATAGTATTTGGGCAAAACGTTGTCGATGACAGG 4724
DB 2294 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTCAATTCAGCTCCGCTGCCCAACGATC 2353
QY 4725 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTCAATTCAGCTCCGCTGCCCAACGATC 4784
DB 2354 AAGCGAGTTATATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 2413
QY 4785 AAGCGAGTTATATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 4844
DB 2414 GATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTCAATTCAGCTCCGCTGCCCAACGATC 2473
QY 4845 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTCAATTCAGCTCCGCTGCCCAACGATC 4904
DB 2474 TAATTCCTTACTGCTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 2533
QY 4905 TAATTCCTTACTGCTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 4964
DB 2534 CAATTCCTTACTGCTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 2593
QY 4965 CAATTCCTTACTGCTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 5024
DB 2594 GATATATACGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTATTGGAACACCTTCTTC 2653
QY 5025 GATATATACGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTATTGGAACACCTTCTTC 5084
DB 2654 GGCGGCAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 2713
QY 5085 GGCGGCAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 5144
DB 2714 TGCACCAAACTTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 2773
QY 5145 TGCACCAAACTTATGATGCTCCCATGTTGTGGCAAAAAGCGTTAGCTCCTTCGCTGCTCC 5204
DB 2774 AGGAAGCAAAATGCGCGCAAAAGGAATAAGGGCGACACCGGAATGTTGAATACTCAT 2833
QY 5205 AGGAAGCAAAATGCGCGCAAAAGGAATAAGGGCGACACCGGAATGTTGAATACTCAT 5264
DB 2834 ACTCTTCCTTTTCAATATTATGAGCATTTATCAGGCTTATGCTCATGAGCGGATA 2893
QY 5265 ACTCTTCCTTTTCAATATTATGAGCATTTATCAGGCTTATGCTCATGAGCGGATA 5324
DB 2894 CATATTGATGATTTAGAAAAATAACAAATAGGGTTCGCGCACATTTCCCGGAAA 2953
QY 5325 CATATTGATGATTTAGAAAAATAACAAATAGGGTTCGCGCACATTTCCCGGAAA 5384
DB 2954 AGTGGCAC 2961
QY 5385 AGTGGCAC 5392

RESULT 13 SYNRLSMV 2964 bp DNA circular SYN 04-JUN-1993
LOCUS Bluescribe KS Minus cloning vector.
DEFINITION Bluescribe KS Minus cloning vector.
ACCESSION L08784
NID 9310728
KEYWORDS Synthetic construct DNA.
SOURCE Synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert, W.
TITLE Obtained from VecBase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program.
ENTRY BLUEKSM
ENTRY BLUEKSM #TYPE DNA CIRCULAR
TITLE Bluescribe KS Minus - Cloning vector
DATE 28-JAN-1987
#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987

```

```

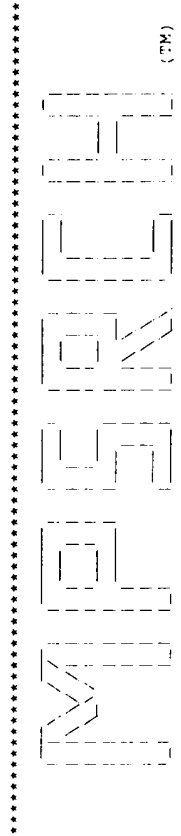
ACCESSION VB0077
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Soroe
J.
#journal Gene (1987) in press
#citation Sequence data from StrataGene
#comment sequence correction according to StrataGene COMMENT
Obtained from StrataGene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer.
1409/10 'AT' to 'TA' to match revised sequence of pBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of StrataGene
Polylinker region revised 03-APR-1987 according to StrataGene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mpl9 lacZ region
is on the complementary strand.
COMMENT
This vector contains the f1 origin so that the minus strand
can be obtained upon f1 superinfection.
KEYWORDS
CROSSREFERENCE
#parent
#parent
VecBase(3):BlueM13m
#parent
VecBase(3):pUC19, VecSource(3):Prom17, VecSource(3):PromT3,
VecSource(3):Bgalks, GenBank(50):PF1
#brother
VecBase(3):BlueKSp, VecBase(3):BlueSKm
PARENT
Features of BlueKSm (2964 bp)
residue source
3- 458 5488-5943 phage f1
460- 524 236- 400 pUC19c
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-2964 448-2617 pUC19c
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueKSm (2964 bp)
residue source
3- 458 5488-5943 phage f1
460- 524 449- 285 (c) pUC19
626- 645 1- 20 T7 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-1031 237- 1 (c) pUC19
1032-2964 2686- 754 (c) pUC19
Conflict (cfl) and Mutations (mut): none
FEATURE
643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-DraII-ApaI-XhoI-SalI-Clal-HindIII-EcoRV-EcoRI-PstI-
SmaI-BamHI-SpeI-XbaI-NotI-XmaII-HstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSm #length 2964 #checksum 1589.
Location/Qualifiers
1..2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 708 a 756 c 735 g 765 t
ORIGIN
source
FEATURES
Query Match 41.2%; Score 2224; DB 32; Length 2964;
Best Local Similarity 99.9%; Pred No 0 006+00;
Matches 2226, Conservative 3, Mismatches 2, Indels 0, Gaps 0;

```


QY 4071 CGGTACACATAGAGGACAGATTTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGG 4130
 Db 6213 AAAAAGAGTTGTAGTCTCTTATCCGCGCAACAAACACGCGTGTGATACGCGTGTGTTT 6272
 QY 4331 AAAAAGAGTTGGTAGTCTCTTATCCGCGCAACAAACACGCGTGTGATACGCGTGTGTTT 4190
 Db 6273 TGTTCGAAGTACGAGATTTACGCGCGAGAAAAAGAGATCTCAAGAGAGATCCCTTCATCTT 5332
 QY 4191 TGTTCGAAGTACGAGATTTACGCGCGAGAAAAAGAGATCTCAAGAGAGATCCCTTCATCTT 4250
 Db 6333 TTCTACGGGCTGTGAGCTCTAGTGGAGCAAACTCACGTTTAAAGGATTTTGGTCATGAG 6392
 QY 4251 TTCTACGGGCTGTGAGCTCTAGTGGAGCAAACTCACGTTTAAAGGATTTTGGTCATGAG 4310
 Db 6393 ATTATCAAAAGAGATCTTACCTAGATCCCTTTTAAATTTAAATGAAGTTTAAATCAAT 6452
 QY 4311 ATTATCAAAAGAGATCTTACCTAGATCCCTTTTAAATTTAAATGAAGTTTAAATCAAT 4370
 Db 6453 CTAAGTATATAGTAAGTAACTTGTGTCACATAGTGTGCTGACTCCCGTGTGTAGAT 6512
 QY 4371 CTAAGTATATAGTAAGTAACTTGTGTCACATAGTGTGCTGACTCCCGTGTGTAGAT 4430
 Db 6513 TATCTACGAGCTGTGCTATTTGCTTCATCCATAGTGTGCTGACTCCCGTGTGTAGAT 6572
 QY 4431 TATCTACGAGCTGTGCTATTTGCTTCATCCATAGTGTGCTGACTCCCGTGTGTAGAT 4490
 Db 6573 AACTACGATACGGGAGGCTTACCTATCTGGCCCGAGTGTGCAATGATACCGCGAGACCC 6632
 QY 4491 AACTACGATACGGGAGGCTTACCTATCTGGCCCGAGTGTGCAATGATACCGCGAGACCC 4550
 Db 6633 ACGTCCAGCGGCTCCAGATTTATCAGCAATTAACACGAGCGGAGGCGGAGCGGAG 6692
 QY 4551 ACGTCCAGCGGCTCCAGATTTATCAGCAATTAACACGAGCGGAGGCGGAGCGGAGCG 4610
 Db 6693 AAGTGTGCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAG 6752
 QY 4611 AAGTGTGCTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAG 4670
 Db 6753 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCAATTCGTACAGGATCGT 6812
 QY 4671 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACGTTGTGCAATTCGTACAGGATCGT 4730
 Db 6813 GGTGTACGCTCTGCTGTTTGGTATGCTTCAATTCAGCTCCGTTTCCCAACGATCAAGCG 6872
 QY 4731 GGTGTACGCTCTGCTGTTTGGTATGCTTCAATTCAGCTCCGTTTCCCAACGATCAAGCG 4790
 Db 6873 ACTTACATGATCCGCCATGTTGTGCAAAAGCGTTAGTCTCTCGTCCCTCCGATCGT 6932
 QY 4791 ACTTACATGATCCGCCATGTTGTGCAAAAGCGTTAGTCTCTCGTCCCTCCGATCGT 4850
 Db 6933 TGTGAGAAGTAGTTCGCCAGTGTATATCCTATGCTGTTATGGCAGCACTGCATAATTC 6992
 QY 4851 TGTGAGAAGTAGTTCGCCAGTGTATATCCTATGCTGTTATGGCAGCACTGCATAATTC 4910
 Db 6993 TCTTACTGTCTATCCATCCGCTAGATGCTTTTCTGTGATGTTGATGATCTACCAACGATC 7052
 QY 4911 TCTTACTGTCTATCCATCCGCTAGATGCTTTTCTGTGATGTTGATGATCTACCAACGATC 4970
 Db 7053 ATTCTGAGAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 7112
 QY 4971 ATTCTGAGAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 5030
 Db 7113 TACCGGCGCATAGTACAGACTTTAAAGTGTCTATCATGTTGGAAGAGTCTTTCGGGGGG 7172
 QY 5031 TACCGGCGCATAGTACAGACTTTAAAGTGTCTATCATGTTGGAAGAGTCTTTCGGGGGG 5090
 Db 7173 AAAACTCTCAGGATCTTACCGCTTGTGAGATCCAGTTCGATGTAAACCACTCGTGACCC 7232
 QY 5091 AAAACTCTCAGGATCTTACCGCTTGTGAGATCCAGTTCGATGTAAACCACTCGTGACCC 5150
 Db 7233 CAACTGATCTTTCAGCATCTTTTACTTTCCAGCGGTTTCTGGTGAGCAAAACAGGAAG 7292
 QY 5151 CAACTGATCTTTCAGCATCTTTTACTTTCCAGCGGTTTCTGGTGAGCAAAACAGGAAG 5210

Db 7293 GCAAAATGCCGAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTACTACTCTT 7352
 QY 5211 GCAAAATGCCGAAAAAGGGAATTAAGGCGACACGGAATGTTGAATACTACTACTCTT 5270
 Db 7353 CCTTTTCAATATTTAAGCAATTTATCAGGTTTATTTGCTCATGAGCGGATACATATT 7412
 QY 5271 CCTTTTCAATATTTAAGCAATTTATCAGGTTTATTTGCTCATGAGCGGATACATATT 5330
 Db 7413 TGAATGATTTAGAAAAATAAACAATAAAGGTTTCCGCGGACATTTCCGCGAAAAAGTCC 7472
 QY 5331 TGAATGATTTAGAAAAATAAACAATAAAGGTTTCCGCGGACATTTCCGCGAAAAAGTCC 5390
 Db 7473 AC 7474
 QY 5391 AC 5392

Search completed: Mon Mar 15 23:40:40 1999
 Job time : 13760 secs.



Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_n n.a. n.a database search using Smith-Waterman algorithm
Run on Tue Mar 16 02:50:01 1999. Maspar time 759.81 seconds
965.496 Million cell updates/sec
Tabular output not generated
Title: >US-09-020-716-6
Description: (1-5192) from US09020716 seq
Perfect Score: 5392
N.A. Sequence: 1 GTAAATTGTAAGCGTTAATA ATTTCGCCGAAAGTGCAC 5392
Comp: GATTTAACATTCGCAATAT TAAAGGCGCTTTCACGGTG
Scoring table: TABLE default
Gap 5
Nmatch STD : DBase 0: Query 0
Searched: 189442 seqs, 58026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: n-genes32
1:part1 2:part3 3:part4 4:part5 5:part6 6:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40
Statistics: Mean 10.519; Variance 7.117; scale 1.478
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2222	41.2	5178	27	T49876	pTet-Splice.	0.00e+00
2	2220	41.2	6206	27	T49877	Autoregulatory vector	0.00e+00
3	2222	41.2	7474	31	T71320	Plasmid pPB73 encodin	0.00e+00
4	2203	40.8	5314	2	Q13576	Plasmid pMTV1 contain	0.00e+00
5	2201	40.8	3681	2	Q13578	Plasmid pKSEL5.	0.00e+00
6	2198	40.8	3699	39	V14340	Plasmid pBSGFP expres	0.00e+00
7	2202	40.8	4145	7	Q40279	Sequence of clone pS2	0.00e+00
8	2198	40.8	4277	7	Q40280	Sequence of clone pS2	0.00e+00
9	2200	40.8	4732	15	Q84596	Plasmid GS control. N.	0.00e+00
10	2200	40.8	5042	15	Q84594	Plasmid glucoamylase	0.00e+00
11	2202	40.8	5356	25	T43794	Plasmid pRIPHAT (rat	0.00e+00
12	2202	40.8	5534	25	T43137	PUMIGIT sequence incl	0.00e+00
13	2199	40.8	5926	7	Q40291	Sequence of clone pP2	0.00e+00

C 14 2198 40.8 7287 39 V02042
15 2196 40.7 4539 14 O87347
16 2196 40.7 10930 20 O81226
17 2196 40.7 10930 20 O81226
18 2195 40.7 14311 24 T38744
19 2190 40.6 6824 5 Q39050
20 2185 40.5 7228 9 Q50201
21 2165 40.2 4118 35 T69189
22 2166 40.2 4118 35 T69189
23 2165 40.2 5585 14 Q87438
24 2165 40.2 6295 39 V02043
25 2162 40.1 4454 30 T56452
26 2162 40.1 4540 32 T69891
27 2162 40.1 4824 20 T88975
28 2163 40.1 4883 4 Q25709
29 2163 40.1 4883 8 Q51192
30 2162 40.1 5399 4 Q25706
31 2163 40.1 5585 29 T51430
32 2163 40.1 5620 4 Q27489
33 2163 40.1 5620 9 Q53875
34 2163 40.1 5642 8 Q51193
35 2162 40.1 5674 7 Q43812
36 2162 40.1 5555 9 Q53874
37 2162 40.1 6266 18 T12652
38 2163 40.1 7050 7 Q40419
39 2162 40.1 7106 8 Q47193
40 2162 40.1 7106 15 Q47494
41 2162 40.1 7298 8 Q47192
42 2163 40.1 10288 31 T71322
43 2162 40.1 10443 31 T71319
44 2163 40.1 13414 31 T71321
45 2162 40.1 14656 21 T30787

ALIGNMENTS

RESULT 1
ID T49876 standard: DNA: 5178 BP.
AC T49876:
DT 28-MAR-1997 (first entry)
DE pTet-Splice.
KW pTet-Splice; pTet-tTAk; tetracycline transactivator, promoter;
KW gene expression; transgenic animal; animal model; drug screening;
KW vector, ds.
FS Synthetic.
FH Key
FT misc_rna
FT complement (2591)
FT Location/Qualifiers
FT /note= "putative start site of trxn"
FT /tag= a
FT complement (2516..2522)
FT /tag= b

WO9640946-A1.
PN 19-DEC-1996.
PD 07-JUN-1996: U010109
PF 07-JUN-1995: US-474169.
PR (UYIA) UNIV YALE.
PI Schatz DG:
DR WPI: 97-077273/07.
PT Nucleic acid encoding tetracycline transactivator fusion protein -
PT provides rapid and reversible control of gene expression, e.g. for
PT creating animal models for drug screening
PS Example 1, Fig 9b-g; 82pp; English.
CC Vector pTet-Splice (T49876) is utilised in the construction of
CC autoregulatory vector plasmid pTet-tTAK (T49877). Splice-PA was
CC made by ligating the SV40 small t antigen intervening sequence and
CC the SV40 early polyA sequence into pBSK1+. The XhoI-SalI fragment
CC of pUHI3.3 (contg. 7 copies of the tet operator upstream of minimal
CC promoter tetp) was cloned upstream of the splice/polyA sequence of
CC pSplice-PA to form pTet-Splice. A modified tetracycline
CC transactivator (tTAK) gene was cloned into pTet-Splice to form pTet-
CC tTAK. The construct provides rapid, reversible control of gene
CC expression in eukaryotic cells of transgenic animals, e.g. for
CC heterologous protein, RNA or antisense sequence production, or for

Db	5057	cccttttcaattatttgaagcatttaccgggtattgtctcatgagcggtacatatatt	5116
QY	5271	CTCTTTTCATATTTTGAAGCAITTTACAGGGTATTGTCTCATGAGCGGATACATAIT	5330
Db	5117	tgaattatttgaataataaacaatagggtttccgcgcacatttcccgcgaagtgc	5176
QY	5331	TGATGTATTAGAAAAATAACAATAAGAGGTTCTCCGCCACATTTCCCGAAAAAGTGC	5390
Db	5177	ac 5178	
QY	5391	AC 5392	
RESULT 2			
ID	T49877 standard; DNA: 6206 BP.		
AC	T49877:		
DT	28-MAR-1997 (first entry)		
DE	Autoregulatory vector plasmid pTet-trak.		
KW	pTet-trak; pTet-Splice; tetracycline transactivator; promoter;		
KW	gene expression; transgenic animal; animal model; drug screening;		
KW	vector; ds.		
OS	Synthetic.		
FS	Location/Qualifiers		
FT	key	complement (2466..3476)	
FT	cds	/*tag= a	
FT	WQ9540946-A1.		
PN	P-PSDB: WQ8474.		
PD	19-DEC-1996.		
PF	07-JUN-1996: U10109.		
PR	07-JUN-1995: US-474169.		
PA	(JYVA) UNIV YALE.		
P1	Schatz DG;		
CP	WPI: 97-077273/07		
DR	P-PSDB: WQ8474.		
PT	Nucleic acid encoding tetracycline transactivator fusion protein -		
PT	provides rapid and reversible control of gene expression, e.g. for		
PT	creating animal models for drug screening		
PS	Example 1: Fig 10b-g; 82pp; English.		
CC	Autoregulatory plasmid pTet-trak (T49877) contains a modified		
CC	tetracycline transactivator gene, trak, under control of minimal		
CC	promoter, Tetp. It was obtd. by cloning the trak gene into vector		
CC	pTet-Splice (T49876). Tetracycline prevents trkA from binding to		
CC	Tetp. preventing expression of trkA and luciferase reporter gene		
CC	When tetracycline is removed, tiny amounts of trkA protein (which		
CC	may result from leakiness of tetp) bind to tet-op and stimulate		
CC	expression of the trkA gene. A positive feedforward loop is		
CC	initiated leading to higher levels of trkA and luciferase. The		
CC	autoregulatory control system allows rapid, reversible control of		
CC	gene expression in eukaryotic cells or transgenic animals, e.g. for		
CC	heterologous protein, RNA or antisense sequence prodn., or for		
CC	creating animal models of human disease.		
SQ	Sequence	6206 BP: 1620 A: 1443 C: 1486 G: 1657 T:	
Query Match 41.2%; Score 2220; DB 27; Length 6206;			
Best Local Similarity 100.0%; Pred. No. 0.00e+00;			
Matches 2221; Conservative 0; Mismatches 1; Indels 0; Gaps			
Db	3985	ctcagagggggcccggtaccagctttttgttcccttttagtgagggttaattgcgcgctt	4044
QY	3171	CTCAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTACGGGTAAATTCGGCGCTT	3230
Db	4045	qccgtaatcagtcacagctgtttccgtgtggaattgttatccgtccacaattccaca	4104
QY	3231	GGCGTAATCATGTCATAGCTGTTTCCTGTGTGAATTTGTATCCGCTCACAATCCACA	3290
Db	4105	caacatacagccgcgaagcataaagtgtaaagcctggggtgcctaatgagtgcactaact	4164
QY	3291	CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAACT	3350
Db	4165	caacataattgccttgcgctcaactccgcgctttccagtcggagaaacctgtcgatccagct	4224
QY	3351	CATATTAAATTTCGTTGGTGTCACTGCGCCGCTTTCCAGTCGGGAAACTTGTGTGTCGAGCT	3410

Db 5853 tttccgacccgtgcgcttacgggataacgtgtccgccttctcccttcgggaagcggtgc 5912
 QY |||||
 Db 3771 TGTTCGACCCGTCCGCTACCGGATACCTGTCGCCCTTCCTCCCTTCGGGAACGGTGGC 3830
 QY |||||
 Db 5913 gctttctatagctacgctgtagtattatcagttcgtgtatagtgcttcccaagct 5972
 QY |||||
 Db 3831 GCTTCTCATAGCTCAGCGTGTAGTATCTCATCTCGGTGTAGTGTCTGCTCCAAAGCT 3890
 QY |||||
 Db 5973 ggcgtgtgacgaaccccccggttcagccgacccgctgcgcttattccggttaactatcg 6032
 QY |||||
 Db 3891 GGGCTCTGTTGTCAGTAAACCTCCCGGTTCAGGCTCCGACCGCTTCGCCCTTATCCGCTAACTATCG 3950
 QY |||||
 Db 6033 tcttgagtcacaccccggttaagacacgacttatcgccactggtgcagcagccactggttaacag 6092
 QY |||||
 Db 3951 TCTTGAGTCCCAACCCGCTAAGACACGACTTATGCGCCACTGCGCAGCAGCACTGGTAACAG 4010
 QY |||||
 Db 6093 gattacagagcagagtagttaggcggtgctacagagttcttgaagtgtggtccctaacta 6152
 QY |||||
 Db 4011 GATTACGAGCGGAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGTGGCCTAACCTA 4070
 QY |||||
 Db 6153 cggctacactagaagacagtagtttggtagctgctgctgctgctgaagccagttactctcgg 6212
 QY |||||
 Db 4071 CGGCTACACTAGAACGACAGTATTTGGTATCTTCGCTCTGCTGAAGCCAGTTACCTTCGG 4130
 QY |||||
 Db 6213 aaaaagttggttagctcttggtagcgggcaaaaacacccaccgctggtgtagcgtgttttt 6272
 QY |||||
 Db 4131 AAAAAGATTGTTAGTCTTGATTCGGCAACCAACACACCGGTGGTACGGGTGTTTTT 4190
 QY |||||
 Db 6273 tgtttcaagcagcagatcacgcgcagaaaaaaggatctcaagaagatcctttgatctt 6332
 QY |||||
 Db 4191 TGTTCGAGCAGCAGTATACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 4250
 QY |||||
 Db 6333 tctacgggtgtgacgctcagtgaaacgaacacacacacacacacacacacacacacacac 6392
 QY |||||
 Db 4251 TTTACGGGCTCTGAGCTCAGTGGAAACCAACTCAGTTAAGGGATTTTGGTCATGAG 4310
 QY |||||
 Db 6393 attatcaaaaagatcttcacatagatccttttaattaaaatgaagttaaataaat 6452
 QY |||||
 Db 4311 ATTATCAAAAAGGATCTTCACCTAGATCTCTTTTAAATTAATAATGAAGTTTAAATCAAT 4370
 QY |||||
 Db 6453 ctatgtatatagataaactgtctacagttacacatgcttaactcaagtgaagcacc 6512
 QY |||||
 Db 4371 CTAAGTATATATAGTAACCTTGTCTGACAGTTTACCAATGCTTTAATCAGTGAGGACC 4430
 QY |||||
 Db 6513 tatctacgcatctgtctatttcttcacatccatagttgctgactcccccgctgtagat 6572
 QY |||||
 Db 4431 TATCTCAGCGATCTGCTATTTCGTTTCATCCATAGTTGCCTGACTCCCGCTCGTGTAGAT 4490
 QY |||||
 Db 6573 aactacgatacggagggttaccatctgcccccgagtgctgcaatgatataccgagagacc 6632
 QY |||||
 Db 4491 AACTACGATACGGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGCGAGACCC 4550
 QY |||||
 Db 6633 acgctacccggtccagatttatcagcaataaaccagcagccggaagggcgagcgagcag 6692
 QY |||||
 Db 4551 ACGCTACCGGCTCCAGATTTATCAGCAATTAACACGCGCGGAGGGCGGAGCGAG 4610
 QY |||||
 Db 6693 aatggtctcgaactttatccgctccatccagctctatattattgttccggggaagctag 6752
 QY |||||
 Db 4611 AAGTGTCTCTGCAACTTTATCGGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAG 4670
 QY |||||
 Db 6753 agtaagtagtccaggttaagtttggcgaacgttggcattgtatacagcattcgt 6812
 QY |||||
 Db 4671 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCGCAACGTTGTGGCAATTTGTACAGCATCGT 4730
 QY |||||
 Db 6813 ggtgtcacgctcgttctgttggtagcttcattcagctccggttcccccaacatcaagcg 6872
 QY |||||
 Db 4731 GGTGTACGCTGCTGTTTGGTATAGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGG 4790
 QY |||||
 Db 6873 agttacatgacccccctgttgtgcaaaaaagcggttagctccttcggtcctccgtagct 6932
 QY |||||
 Db 4791 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTCGTTCCGATCGT 4850
 QY |||||

Db 6933 tdtcagaagtaagtggccgcagtggttatcactcatgttatggcagcactgcataatc 6992
 QY |||||
 Db 4851 TCTCAGAAGTAAGTTGGCGGAGTCTTATCACTCATGTTATGTCAGTATGTCAGTATGTC 4910
 QY |||||
 Db 6993 tcttactgtatgcatacgcgttaagatgcttttctgtgtgactggtgagtaactcaaccagtc 7052
 QY |||||
 Db 4911 TCTTACTGTATGTCATCCGTAAGAAGTCTTCTGTACTGTGAGTACTCAACCAAGTC 4970
 QY |||||
 Db 7053 attctgaaatagttgtgacgcacgagttgcttcttgcggcgtcaatacgggataa 7112
 QY |||||
 Db 4971 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGCGTCAATACCGGATAA 5030
 QY |||||
 Db 7113 taccgcgcacatacagcaaacctttaaaagtgctcatcattggaacacgttcttcggugcg 7172
 QY |||||
 Db 5031 TACCGCGCACATACAGCAACTTTTAAAGTGCTCATCATTTGGAAAACTGTTCTTGGAGCG 5190
 QY |||||
 Db 7173 aaactctcaaggtacttaccgctggttgagatccagttcgaatgaacccactcgtgcacc 7232
 QY |||||
 Db 5091 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCTGTGACC 5150
 QY |||||
 Db 7233 caactgactctcagcatcttttactttccacgagcttcttgggtgagcaaaaacagaaag 7292
 QY |||||
 Db 5151 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 5210
 QY |||||
 Db 7293 gcaaatgccgcacaaaggaataaaggcgacacacgcaaaatgttgatactcatacctt 7352
 QY |||||
 Db 5211 GCAAAATGCGCAAAAAAGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTT 5270
 QY |||||
 Db 7353 ccttttcaatattattgaagcattatcagggttattgtctctcagcgagataca-att 7412
 QY |||||
 Db 5271 CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAIT 5330
 QY |||||
 Db 7413 tgaatgtattagaataaataaacaataaggggttcgcgcacatttcccgaaaaagtcgc 7472
 QY |||||
 Db 5331 TGAATGTATTAGAAAAATAACAATAAGGGGTTCCCGGCACATTTCCCGGAAAAAGTGC 5390
 QY |||||
 Db 7473 ac 7474
 QY ||
 Db 5391 AC 5392

RESULT 4

ID Q13576 standard; DNA: 5314 BP.
 AC Q13576;
 DT 02-DEC-1991 (first entry)
 DE plasmid pMTV1 containing 3 multiple cloning sites.
 KW recombinant expression plasmid; phage phi X174; protein E;
 KW bacterial ghost; ss.
 OS Synthetic.
 PN W09113155-A.
 PD 05-SEP-1991.
 PF 19-FEB-1991; E00308.
 PR 24-FEB-1990; DE-005874.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Lubitz W, Szostak MP;
 WP1: 91-281471/38.
 PT Carrier recombinant protein for vaccines against HIV, etc.
 PT obtd. by expressing fusion protein gene in gram negative bacteria
 PT and gene that encodes for lytic membrane protein
 PS Example 10; Page 28-29; 45pp; German.
 CC The carrier plasmid pMTV1 comprises the phage phi X174 E protein
 CC gene, the ampicillin resistance gene, the lambda ci857 repressor
 CC gene and P(R) from the lambda promoter/operator system. There are
 CC also lac operon sequences and three multiple cloning sites. A
 CC foreign gene can be inserted into the second multiple cloning site
 CC (mcs2). The protein encoded by the foreign gene is especially a
 CC viral antigenic peptide, e.g. of HIV, HBV or EBV.
 CC See also Q13577 and Q13578.
 SQ Sequence 5314 BP; 1438 A; 1276 C; 1310 G; 1290 T;

Query Match 40.9%; Score 2203; DB 2; Length 5314;

Best Local Similarity 99.5%; Pred.No. 0.00e+00;

Matches 2215; Conservative 0; Mismatches 12; Indels 0; Gaps 0;


```
QY 3911 CCGTTCAGGCGGCGGCTGCGCTTATCGGTAACTATCGTCTAGTCCACCGCGTAA 3970
Db 2275 gacacgacttatcgccactgcacagccactggtaacagagattagcagacgagatag 2334
QY 3971 GACACGACTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGATTAGCAGACGAGGTATG 4030
Db 2335 tagcggtgctacagagttcttgaagtgggtggtggtggtggtggtggtggtggtggt 2394
QY 4031 TAGGGGCTGCTACAGAGTCTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4090
Db 2395 tatttgatctggtgctgctgagccaggttacctcggaagaaagtggtgagctctt 2454
QY 4091 TATTGGTATCTGCGCTCTGCTGAAGCGAGTACCTTCGGGAAAAGAGTTGGTGTGCTCT 4150
Db 2455 gacccgcaaaacacaccccgctgtagcggtggttcttcttcttcttcttcttcttctt 2514
QY 4151 GATCCGGCAACCAACTACCGCTGCTGAGCGGTGTTTTTTTGTGTCAGCAGCAGATTA 4210
Db 2515 cgcgcagaaaaaagagatctcaagaagatcccttctgactcttcttcttcttcttcttctt 2574
QY 4211 CATGAGAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTC 4270
Db 2575 atggaacaaacactcaacttaagagatttctgctcagatattcaaaaaagatcttca 2634
QY 4271 AGTGAACGAAACACTCAGCTTAAAGGATTTTGGTCAATGATATCAAAAAGGATCTTCA 4330
Db 2635 cctgagctctttaaataaataaagattttaaataaataaataaataaataaataaataa 2694
QY 4331 CCGATGCTTTTAAATTAATAATGAAGTTTAAATCAATCAATGAAGTATATGAGTAA 4390
Db 2695 cttgctcagcagttaccactgcttaatacagtgagggcaccctatctcagcagctctat 2754
QY 4391 CTGCTCTCAACAGTTACCAATGCTTAATCACTGAGGACCACTATCTCAGCGATCTGCTAT 4450
Db 2755 tctgctcactcagttctgctgactccctcgtgctgctgctgctgctgctgctgctgct 2814
QY 4451 TCGCTTCATCATAGTTGCTTASATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4510
Db 2815 taccatctggcccccagtgctgcaatgataccgcgcagaccacgcgtccacggctccagatt 2874
QY 4511 TACCATCTGGCCCCAGTCTGCTAATGATACCGCGAGACCACCGCTCACCGGCTCCAGATT 4570
Db 2875 tatagcaataaaccagcagccggaagggccgagcgcagaaagtgtcctgcaactttat 2934
QY 4571 TATCACTAATAAACCCAGCCGAGGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 4630
Db 2935 cgcgcctcactcagctctatattgttgcgggaagctagagtaagtagtgcgcagttta 2994
QY 4631 CGGCTTCATCCAGTCTATTAATGTTGCCGGGAGGCTAGAGTAAGTAGTTCGCCAGTTA 4690
Db 2995 atagttgcgcaacgtttgtgctgctgctgctgctgctgctgctgctgctgctgctgct 3054
QY 4691 ATAGTTTGGGCAACGCTTCTGCTCAATGCTAGTCAATGCTGCTGCTGCTGCTGCTGCT 4750
Db 3055 gtagcttcactcagctccggttcccaacgatacagagcagattacatgatacctccatgt 3114
QY 4751 GTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGT 4810
Db 3115 tctgcaaaaaagcgttagctcctcgtcctcgtcctcgtcctcgtcctcgtcctcgtcctcgt 3174
QY 4811 TGTGCAAAAAGCGGTTAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4870
Db 3175 cagttattcactcagttatggcagcactgcataatctcttactgtcagtcagtcagtcagtc 3234
QY 4871 CAGTGTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACGTGCTATGCCATCCG 4930
Db 3235 taagatcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3294
QY 4931 TAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGC 4990
Db 3295 ggcgacccagttgctctgcccgcgcgtcaatacagggataataacgcgcacacatagcaga 3354
```

```
QY 4991 GGCACCGAGTGTGCTCTTGCCGGGCTTAATAAGGATAAIAAGGAGACALAGCAGAA 5050
Db 3355 ctttaaaagtgcctcatttgaaacacgttcttcgggggcaaaactctcaagatcttac 3414
QY 5051 CTTTAAAGTGTCTCATCTTGSAAACGTTCTTTCGGGGGCAAACTCTCAAGGATCTTAC 5110
Db 3415 cgtgttgagatccagttcagatgaacccactcgtgcacccactgactcttcagatctt 3474
QY 5111 CGCTGTTGAGATCCAGTTCGATGTACCCACTCGTGCACCCCACTGATCTTCAGCATCTT 5170
Db 3475 ttacttccacagcgtttctgggtgagcaaaaaacaggaagcaaaatgccgcaaaaaag 3534
QY 5171 TTACTTTTCCAGCAGGTTCTGCTGGTGAGCAAAAAAGAGAAATGCGCAAAAAGAG 5230
Db 3535 gaataagggcagacagaaatgttgaatactcactcactcactcactcactcactcactcact 3594
QY 5231 GAATAAGGGGAGACCGGAAATGTTGAATACTCATCTCTCTCTCTCTCTCTCTCTCTCT 5290
Db 3595 gcatttatcagggttatgtctcatdaggggatacatatttgaatgatttadaaaaata 3654
QY 5291 GCATTTATCAGGTTATTGTCTCATGAGCGGATACATATTGAATGATTATTAGAAAAATA 5350
Db 3655 acaaataggggttcgcgcacatttcccccgaagaagtgccac 3696
QY 5351 AACAAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGTCAC 5392
```

RESULT 7

```
ID Q40279 standard: DNA: 4145 BP.
AC Q40279;
DE 02-AUG-1993 (first entry)
DE Sequence of clone pS2gpt-S4.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..2226 /*tag= a
FT /*label= pN2gpt-S4
FT /*note= "Posn. 1 corresp. to the first nucleotide
FT G '5-TGCACACTTT TCGGGGAAAT-3'."
FT misc_feature 2227..2236 /*tag= b
FT /*label= SmaI adaptor
FT 2396..2851 /*tag= c
FT /*label= E. coli qpt gene
FT promoter complement (3081..3323)
FT /*tag= d
FT /*label= vaccinia p7.5 promoter
FT promoter 3358..3451 /*tag= e
FT /*label= S4 of pN2gpt-S4
FT /*note= "oligo P-artP(9)."
FT misc_feature 2237..4145 /*tag= f
FT /*label= pN2gpt-S4
```

```
PN AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO ) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
DR WPI: 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example: Pages 160-162; 206pp; English.
CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli qpt
CC gene driven by the vaccinia virus p7.5 promoter, flanked by
CC several unique restriction sites including NotI. The plasmid
```


QY 5271 CTTTTTCAATATTATGAAGCATTATATCAGGCTTATTTGCTCATGAGCGATACATATT 5330
Db 4978 tgaatgtatttgaataataacaaatagggtttccgcacatttcccgaaagtgc 5037
QY 5331 TGAATGATTATGAAAAATAAACAATAAGGGTTTCCGGCACATTTTCCCGGAAAGTGCC 5390
Db 5038 ac 5039
QY 5391 AC 5392

RESULT 11
ID T43794 standard; DNA: 5356 BP.
AC T43794:
DI 18-FEB-1997 (first entry)
DE plasmid PRIPHAT (rat insulin promoter-human IAPP transgene).
KW type II diabetes mellitus; transgenic animal model; pancreas;
KW islet; beta cell; islet amyloid polypeptide; IAPP; hyperglycaemia;
KW glycaemia; diabetic glomerulosclerosis; plasmid PRIPHAT; ds;
KW cyclic.
OS Chimeric Homo sapiens;
OS Chimeric Rattus sp.;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FI promoter 719..1594
FI /tag= a
FI /label= RIP-II-promoter
FI /note= "rat insulin II promoter"
FI cds 1589..1866
FI /tag= b
FI /label= IAPP
FI /note= "human IAPP coding sequence (Claim 7)"
FI intron 1861..2580
FI /tag= c
FI /label= Albumin.intron
FI /note= "human albumin intron I"
FI terminator 2575..3119
FI /tag= d
FI /label= Poly-A
FI /note= "human GAPDH gene poly-A and RNA
FI termination region"
PN WO9637612-A1.
PD 28-NOV-1996. IB0371.
PF 24-APR-1996. IB0371.
PR 23-MAY-1995. US-446935.
PA (F12) PFIZER INC.
PI Carty MD, Kreutter DK, Soeller WC;
PI WPI: 97-021221/02.
PT Recombinant DNA for expression of islet amyloid polypeptide - to
PT develop prods. for use in diagnosis, study and treatment of
PT disorders, e.g. diabetes and obesity
PS Claim 6: Page 21-23; 46pp; English.
CC Plasmid PRIPHAT (T43794) allows pancreatic beta cell-specific
CC expression of human islet amyloid-associated polypeptide (IAPP)
CC under the regulation of the rat insulin II promoter in both cell
CC lines and in transgenic animals (models of type II diabetes
CC mellitus). It was constructed by PCR amplification (see also
CC T43795-804) of the relevant DNA fragments and their insertion into
CC a pBluescript SK(-) framework. The transgenic animals develop
CC amyloid plaque deposits in islet cells, fasting hyperglycaemia,
CC glycaemia and diabetic glomerulosclerosis. They can be screened
CC for treatments that modulate disease progression. Cell lines can
CC be screened for treatments that inhibit human IAPP expression.
SQ Sequence 5356 BP; 1368 A; 1294 C; 1274 G; 1420 T;
Query Match: 40.8%; Score 2202; DB 25; Length 5356;
Best Local Similarity 100.0%; Pred No. 0.00e+00;
Matches 2202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3155 ccagctttgttcccttttagtgaggttaattgcgcgttgcgttaattcgttcacggttcgaac 3214
QY 3191 CCAGCTTTTGTCCCTTTAGTGAGGTTAATTGCGCGCTTGGCGTAAATCATGTCATAGC 3250

Db 3215 tgttctgtgtgaattgttatccgctcacaattccacacacatacagcgccggaac 3274
QY 3251 TGTTCCTGTGTGAATTTGTATCCGCTCACAATTCACACACATACGAGGCGGAGGCA 3310
Db 3275 taaagtgtaaagcctgggtggtcetaatagtgagctaaetccacatttaattggttgcgt 3334
QY 3311 TAAAGTGTAAAGCCTGGGCTGCTTAATGAGTGAAGTAACTCACATTAAATTCGCTGGCT 3370
Db 3335 cactgcgcgtttccagtcgggaaacotcgtgcgaactccatcaatcaatcggcgaac 3394
QY 3371 CACTGCCGCTTTCAGTTCGGGAACCTGCTGCTAGCTGCAATTAATGAATCGGCAAC 3430
Db 3395 gcgcgggagagcggtttgcgtattggcgctcttccgcttccctccctccacacactcgc 3454
QY 3431 GCGGCGGAGAGGCGGTTTGGGTATTTGGGCGCTTTCGCTTCCTGCTACATGACTCG 3490
Db 3455 tgcctcgttgcgttcgctgcggcgagctatcagctcactcaaaagcggttaatacgggt 3514
QY 3491 TCGGCTCGGCTTTCGCGTGGGAGCGGTGCTGCTGCTTTTCCATAGGCTTCGCGCTCAG 3550
Db 3515 tatccacagaatcaggggataaacccaggaagaacatgtgacaaaggccacaaagg 3574
QY 3551 TATCCACAGAAATCAGGGGATTAACCCAGGAAGAACATGTGAGCAAAAGGCCACCAAGG 3610
Db 3575 ccaggaaacctaaaggccgcgttgcgtgctgttttccataggtccgcgcgcctcgaag 3634
QY 3611 CCAGGAACCGTAAAAAGCGCGCTTGTGCTGCTTTTCCATAGGCTTCGCGCTCAGG 3670
Db 3635 agcatcacaataatcgagcgtccaagtgcaggtggcgaaaccccgacagactataagat 3694
QY 3671 AGCATCACAATAATCGAGCGCTCAAGTCAGAGGTGGCGAAACCCGACAGSACTATAAGAT 3730
Db 3695 accagcgcttcccccctggagctccctcgtgcgctcctcgttcgcgaacctgcgctta 3754
QY 3731 ACCAGCGCTTTCCCGCTGGAGAGTCCCTCGTGGCTCTCTCTGTTCCGACCTGCCGCTTA 3790
Db 3755 ccggatacctgcgccttttcccttcgggaagcggtggcgctttctcatagctcacgt 3814
QY 3791 CCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCT 3850
Db 3815 gtaggatctcagtcgggtgagtccttcgctccaaagctggctgtgtgcagaacccc 3874
QY 3851 GTAGGTATCTCAGTTCCGTTGAGTGTCTGCTCCAGCTGGGCTGTGTGCACGAGACCC 3910
Db 3875 ccgttcagcccgacgcgtgcgccttaaccgtaactcctcttgatcccaaccccgataa 3934
QY 3911 CCGTTTCAGCGCGACCGCTGCGCTTATCCGCTTAATCTGCTTGTAGTCCACCCGCTAA 3970
Db 3935 gacagacttatccgactgcgcagccactgtaacaggtatgagagagaggtatg 3994
QY 3971 GACAGACTTATCCGCTTATGCTAGCAGTCACTGTAAACAGSATTAGTAAAGTGAATAG 4030
Db 3995 tagcggtgctcacaggtcttgaagtgggtggccctaactacgactacactagaagacag 4054
QY 4031 TAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGAGACAG 4090
Db 4055 tatttggtatctgcgctcgtcgtgaagccagttaccttcgaaaaagagttggtagctctt 4114
QY 4091 TATTGTTATCTGCTGCTGCTGAAGCCAGTAACTTCGAAAAAGAGTGTGTAGTCTTT 4150
Db 4115 gatcgggcaacaaacacacgcgtgtagcggtggttttttgggttgcgaacacacaaatta 4174
QY 4151 GATCGGGAACAAACACCGCTGCTAGCGGTGCTTTTGTGTTGGAAAGCAGCATTA 4210
Db 4175 ccgcgagaaaaaaggatctcaagaagatccttgcattcttccagggctcgaacctc 4234
QY 4211 CCGCGAGAAAAAAGATCTCAAGAAATCCTTTTATCTTTTCTACGCGGTCTGAGCGCTC 4270
Db 4235 agtgaacgaaaaactcagtttaagggtatttggctcatgagattatcaaaaagatcttcca 4294
QY 4271 AGTGAACGAAAAACTCAGCTTAAGGGATTITGGTCATGATATATCAAAAGAGATCTCA 4330
Db 4295 cctaagatccttttaataataaataagagttttaaatcaactcaatgatatatgagtaaa 4354

[illegible]

ID	T43137 standard; DNA: 5534 BP.
AC	T43137:
DG	10-FEB-1997 (first entry)
DE	pMIGIT sequence including upstream activation sequence,
KW	Transgenic plant; gene expression; upstream activating sequences;
KW	OAS; transactivating protein; Gal4; herbicide resistance;
OS	Polyhydroxybutyrate; safety; pMIGIT ds.
SQ	Synthetic.
CN	CAZ150039-A.
PD	09-AUG-1996.
PF	24-MAY-1995; 150039.
PP	(X)-FEB-1995; GR-002456.
PA	(UYWA-) UNIV WARWICK.
PI	Hennett M., May S., Ramsay N.
PR	WPI: 96-486150/49.
PT	Control of genes in transgenic plants - using an upstream activating
PT	sequence activated by a transactivating protein expressed using a
PT	separate promoter
CC	Example: Fig 5f: 48pp; English.
CC	Reporter plasmid pMIGIT (143137), or pHAS Minimal Promoter-Gus
CCC	Int Terminator, contains a beta-glucuronidase (GUS) reporter
CCC	gene under the control of the 4kb cauliflower mosaic virus minimal
CCC	promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream
CCC	activation sequence (OAS) of pMIGIT is activatable by yeast
CCC	transactivating protein GAL4. In a novel method for controlling
CCC	gene expression, a first transgenic plant carrying a gene encoding
CCC	a desired phenotype (herbicide resistance of polyhydroxybutyrate
CCC	prodn) operatively linked to a OAS recognition site is pollinated
CCC	by a second transgenic plant carrying Gal4 DNA (see also 143146).
CCC	The transgene is fully expressed in F ₂ -hybrid plants but segregates
CCC	apart in subsequent generations, improving environmental safety.
CC	Sequence 5534 BP: 1400 A; 1385 C; 1487 G; 1403 T;
SQ	
<hr/>	
Query Match 40.8% Score 2202 DB 25 Length 5534	
Best Local Similarity 100.0% Pred No. 0 00e+00;	
Matches 2302 Conservative 0 Matches 0 Indels 0 Gaps 0	
DB	3329 atgacacttttctgggaatatacagcaaacccattgttatatttttaaaatat 4099
Gp	
Gp	5342 GTGGTACTTTTGAGGAAGAATCTGTGTAAGTCATACTGTATTATTTCAAATAAAT 5434
Gt	
Gt	3389 caaatgatgcacctcatcgpaatagaatcataatcatttcagaaaatcccacaan 4459
Gt	
Gt	5332 CAACAATCGATCAGCATGCAGAACAAAGAAAAGAAAAAGAAAAGAAAAGAAA 5274
Db	
Db	3449 gaaagagttagtgattcacacatttccttgtgccctattaccattttttgcacatttt 4508
Gp	
Gp	5272 GGACAGATACAGTAGTTCACATCTGGTGTGTCGCCCTTATTGCCCCCTTTCGGCATTTG 5214
Db	
Db	3509 gcccttcgttttttgtctcaccaccaaaaaacacatcatgaadacaaatata 4648
Gp	
Gp	5215 GCATGCTGTGTTTTTCTGTATGTAAGAAAAGTGTGTAAAATAGTGTGAATATGACT 5153
Db	
Db	3569 tggctgcacqagtgggttacatccaactgaactcaacaggatgaataccttcaaagt 4628
Gp	
Gp	5152 TGCGTGTAAGTACTGGATACATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5094
Gt	
Gt	3620 ATTGCGCGCGAGAGAGCTTTTTCGAATGATGATGATGATGATGATGATGATGAT 4688
Gp	
Gp	5092 TTGCGCGCGAGAGAGATTTTTCGAATGATGATGATGATGATGATGATGATGATGAT 5046
Db	
Db	3689 tatataccccgattgagaccgagcaaanataataatgaacacatcctaacttctcaga 4746
Gp	
Gp	5032 TATTATGCGCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4974
Db	
Db	3749 atgaccttggttgadtactcccaactcaccadaaadacattcaagaataatgaatana 4898
Gp	
Gp	4972 ATGACTTGCTGCTACTCACACCAGTCACAGAAAGATCTTAGCGCATGCAATGATAA 4914
Db	
Db	3809 gaqaatatctycagtgacttctctatdaacatalgaattgaatactatgaacattgattcga 4964
Gp	
Gp	4912 GAQAATTATGCAGTGCTGCTCATAGCTATAGATGATGATGATGATGATGATGATGATGAT 4862

Db 3869 caacatcgagagcgaagagcgaacccgcttttttgcaacacatgggggacatgtaa 3928
Cp 4852 CAAGATCGAGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGATCATGTAA 4793
Db 3929 ctgcgcttgatggttgggaacccgagctgaatgaagccatccaaacagcagcgtgaca 3988
Cp 4792 CTGCGCTTTGATGCTTGGAAACGCGAGCTGAATGAAGCTATACCAACGACGAGCGTGACA 4733
Db 3989 ccacatgctgcttagcaatggcaacacgcttgcgcaaaacttaactggcgaaactactta 4048
Cp 4732 CCACGATGCTGTAGCAATGSCACAAAGCTTCCGCAAACTATTAACTGGGAACTACTTA 4673
Db 4049 ctctagcttcccggcgaacaataatagactgagatgagcgagataaagtgcagggaccac 4108
Cp 4672 CTCIAGCTTCCGCGGCAACAAITTAATAGACTGATGGAGCGGATAAAAGTTGCAGGACCAC 4613
Db 4109 ttctacgctcgaccccttcggcgtgctgctgcttattgtgtgataaactcggagcggtagc 4168
Cp 4612 TTCGTGGCTCGGCGCTTCGGGCTGCTGGTTATTTGCTGATAAATCTGGAGCGGCTGAGC 4553
Db 4169 gtgggtctgcggtatcatctgcagcactggggccagatggtaagccctcccgatcgtag 4228
Cp 4552 GTGGGTCTCGCGGTATCATTTGCAGCACTGGGCGCAGATGGTAAGCCCTCCCGTATCGTAG 4493
Db 4229 ttatctacacgacggggagtcaggcaactatggatgaacgaataatagacagatcgctgaga 4288
Cp 4492 TTATCTACACGACGGGAGTCAAGCAACTATGGATGAACGGAATAGACATCGCTGAGA 4433
Db 4289 taagtgcctcactgattaaagcattgtaactgctcagaccagatttaactatatacttt 4348
Cp 4432 TAGTGCCCTCACTGATTAAAGCAATTGGTAACGTGTACAGCAAGTTTACTCATATATACTTT 4373
Db 4349 aaattaatttaaaacttcatttttaataaaagatctagtgtaagatccctttttgata 4408
Cp 4372 AGATTGATTTAAACTTCATTTTAAATTTAAAGGAICTAGTGAAGATCCCTTTTGATA 4313
Db 4409 atctcatgacaaaatcccttaactgaactgaacttttgcgtccactgagcgtcgagcccttag 4468
Cp 4312 ATCTCATGACCAAAATCCGCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAAGCCCCGTAG 4253
Db 4469 aaaaatcaaaagattctttgagatccctttttctgcgcgtaactctgctgctgcaaa 4528
Cp 4252 AAAAGATCAAGGATCTCTTGAGATCCTTTTTCGCGCGTAATCTCTGCTTGGCAA 4193
Db 4529 caaaaaaccacccgtaccagcgggtggtttgttgcgggatacaagagctaccaaactttt 4588
Cp 4192 CAAAAAACAACCGCTACCAAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTT 4133
Db 4589 ttccgaagtaactggcttcagcagcagcagacagataccaaataactgctctcttagtagc 4648
Cp 4132 TTCGGAAGGTAACTGGCTTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGC 4073
Db 4649 cgtagttagccaccacttcaagaactctgtagcagccgctacatactcgtctgtaa 4708
Cp 4072 GGTAGTTAGGCCACACITCAAGAACTCTGTAGCACCGCTTACATACCTCGCTGTGTAA 4013
Db 4709 tctgttaccagtggctgctgcagcagcagataagctgcttaccggggttggaactcaa 4768
Cp 4012 TCTGTTTACAGTGGCTGTGCTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTGACTCAA 3953
Db 4769 gacatagttagcagataaggcgcagcgtcgggctgaacgggggttcgtgcacagc 4828
Cp 3952 GAGCATAGTTACCGTATAGGCGCGAGCGCTCGGCTGAACCGGGGGTTTCGTGACACAGC 3993
Db 4829 ccagcttgagcgaacgacctacacccgaactgagataactacagcgtgagctatgagaaa 4888
Cp 3892 CCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCIACAGCGTGAAGTATGAGAAA 3833
Db 4889 ggcgcagcttcccgaaggagaaaggcgacagctatccggtaagcggcagggtcgaaa 4948
Cp 4832 GCCCCAAGCTTCCGGAAGGAGAAAGGGCGACAGGATCCGGTAAGCGGACGGGTCGGAA 3773
Db 4949 caggagagcgaacgagggagcttccaggggggaacgcctggtatctttatagtcctgtagc 5008

Cp 3772 CAGGAGAGCGCACGAGGAGGTTCCAGGGGGAACGCTGCTATCTTATAGTCTGTGICG 3713
Db 5009 ggtttcgcaactctgactgagcgtcgatatttggatgactcctcaggggggcaagcc 5068
Cp 3712 GGTTCGCGACCTCTGACTTTCAGCGTCSATTTTGTGATGCTGTCAGGGGGGCGAGGCC 3653
Db 5069 tatggaataacccagcaacgagcgcctttttagcgttccctggccttttgcgtgaccttttg 5128
Cp 3652 TATGAAAAAGCCAGCAACGCGGCTTTTACGGTTTCCTGGGCTTTTGTGGGCTTTTG 3593
Db 5129 ctacatgatttttctcgtgcttattccctctgattctgtggataaccctattaccgcctttg 5188
Cp 3592 CTCACATGTTCTTCCTCGCTTATCCCTCTGATTCTGTGGATAACCGTATTACCGCCTTTG 3533
Db 5189 aatgaactgataaccgctgcgcgaacgcaacgcaacgcaacgcaacgcaacgcaacgcaacg 5248
Cp 3532 AGTGAGCTGTATACCGCTGCGCGTAGCGGAACGAGCGGAGGAGGAGGAGGAGGAGG 3473
Db 5249 aagcggaagagcgcccaataacgcaaacgagcctctcccgcgcttggccgacttcaataa 5308
Cp 3472 AAGCGGAAGAGCGGCCAATACGCAAAACGCGCTCTCCCGCGGCGTGGCGGATTCATTA 3413
Db 5309 gcagctggcagcagaggtttcccgactggaaagcgggcgagtgagcgcaacgcaacgcaacg 5368
Cp 3412 GCAGCTGCGACGACAGGTTTCCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATG 3353
Db 5369 tgagttagctcactcattagcagcccgagccttaccgctttacactttatgcttccggtcgatgt 5428
Cp 3352 TGAGTTAGCTCCTCACTATTAGGCAACCGGCTTTTACACTTTATGCTTCGCGCTCGTATGT 3293
Db 5429 tgtgtggaattgtgagcggaatacaatttcacacagaaacagctatgacactgattacg 5488
Cp 3292 TGTGTGGAATTGTGAGCGGATAACAATTTTCACACAGGAACAGCTATGACCATGATTACG 3233
Db 5489 ccaagcgcaataaacctcactcaataagggaacaaagctga 5530
Cp 3232 CCAAGCGGCAATTAAACCTCACTAAAGGGAACAAAGCTGG 3191

RESULT 13

ID Q40281 standard; DNA; 6926 BP.
AC Q40281.
DT 02-AUG-1993 (first entry)
DE Sequence of clone pP2-gp160MN.
KW Plasmid; cloning; restriction site; HIV-1; MN-isolate; ss.
OS Synthetic.

Key	Location/Qualifiers
misc_feature	1..3529
FT	/*tag= a
FT	/label= pS2gpt-P2 sequences
FT	2396..2851
FT	/*tag= b
FT	/label= E. coli gpt gene
FT	complement (3081..3323)
FT	/*tag= c
FT	/label= vaccinia P7.5 promoter
FT	2396..2851
FT	/*tag= d
FT	/label= P2 according to EP application Avipox
FT	"intergenic region"
FT	/note= "oligo P-artp(9)"
FT	6173..6926
FT	/*tag= e
FT	/label= pN2gpt-S4
FT	3534..6001
FT	/*tag= f
FT	/label= HIV-1 strain MN gp160 sequence (EMBL to
FT	REHIVMNC)

PN A09221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.

PB 20-JUL-1992: US-914738.
 PA (IMMO) IMMUNO AG, Pfeleiderer M, Scheiflinger F;
 PR Werner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
 PP WPI: 93-125451/15.
 PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
 PT direct molecular cloning of modified DNA molecule using
 PT cytoplasmic DNA virus genome
 PS Examples: Pages 177-180, 205pp; English.
 SC Plasmids pN2apt-S3A and pN2apt-S4 comprise expression cassettes
 SC with a selective marker. These plasmids were constructed by first
 SC making plasmids pN2-aptA and pN2-aptB which contain an E. coli qpt
 SC gene driven by the vaccinia virus p7.5 promoter, flanked by
 SC several unique restriction sites including NotI. A new SmaI site
 SC was introduced into the plasmid pN2apt-S3 resulting in the plasmid
 SC pN2apt-S4. Subsequently the S4-promoter was exchanged by the p2-
 SC promoter resulting in the plasmid pN2apt-P2. For the modification
 SC of the qpt60-gene, a PCR-generated proximal fragment was exchanged
 SC leading to a qpt60-gene cassette with a minimal 5'-UTP
 SQ Sequence 5925 BP: 2064 A: 1470 C: 1496 G: 1706 T:

Query Match 40.8% Score 2198; DB 7; Length 5926;
 Best Local Similarity 99.9%; Prod No. 0.00e+00;
 Matches 2280; Conservative 60; Mismatches 23; Indels 0; Gaps 0;

Db 1 atgcaacttttgcggggaatgaggggaacccctattgtttatttttctaaacatt 60
 Cp 5342 GTGGAATTTTGGGGAATGATGGGGGAACCCCTATTGTTTCTTAATACATT 5333
 Db 51 caatatatccctcatataraacataaacctatataatcttaataattatgaataa 120
 Cp 5342 CAATAATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 5333
 Db 121 gaaagatataatgattacacatttttgcgttgccttattccctttttgcgcatitt 180
 Cp 5272 GGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5213
 Db 141 gctttctatttttgcacacacacacacacacacacacacacacacacacacac 240
 Cp 5212 GCTTTCTATTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5153
 Db 241 agaataraagatgattacacattgattcattcattcattcattcattcattcatt 300
 Cp 5152 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5093
 Db 401 ttgcac 360
 Cp 5092 TTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5033
 Db 601 tatataac 420
 Cp 5012 TATTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4973
 Db 141 atgacttgaattgaattgaattgaattgaattgaattgaattgaattgaattga 480
 Cp 4972 ATGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4913
 Db 481 gaaatatacagctgctgctgctgctgctgctgctgctgctgctgctgctgctg 540
 Cp 4912 GAGAATTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4853
 Db 541 caacatatacagctgctgctgctgctgctgctgctgctgctgctgctgctgct 600
 Cp 4852 TAAATGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4793
 Db 601 ctgccttgaattgaattgaattgaattgaattgaattgaattgaattgaattga 660
 Cp 4792 CTCGGCTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4733
 Db 661 caacatatacagctgctgctgctgctgctgctgctgctgctgctgctgctgct 720
 Cp 4712 GCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4673
 Db 721 ctctaatattgcggaacacacacacacacacacacacacacacacacacacac 780

Cp 4672 CTATAGCTTCCCGGCAACAATTAATAGACTGATGAGCGGATAAATCTGACAGATTA 4613
 Db 781 ttctgcgtctgcgcctctgcgcctgcgcctgcgcctgcgcctgcgcctgcgcctgc 840
 Cp 4612 TTCTGCGTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCG 4553
 Db 841 gtaagctctgcgcctctgcgcctctgcgcctctgcgcctctgcgcctctgcgcct 900
 Cp 4552 GTGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCG 4493
 Db 901 ttctctac 960
 Cp 4492 TTATCTACAGCAAGCGGCAATTAATAGACTGATGAGCGGATAAATCTGACAG 4433
 Db 961 taagct 1020
 Cp 4432 TAGGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCG 4373
 Db 1021 aatatttaaacact 1080
 Cp 4372 AGATTGATTTAAACCTTCATTTTAAATTTAAAGGATTAAGCTTAAGCTTAAG 4313
 Db 1081 atctctatgcac 1140
 Cp 4312 ATCTCATGATCAAAATTCATTTAAAGGATTAAGCTTAAGCTTAAGCTTAAG 4253
 Db 1141 aaaaatcaaaaatct 1200
 Cp 4252 AAAAGATCAAAATTCATTTAAAGGATTAAGCTTAAGCTTAAGCTTAAGCTTA 4193
 Db 1201 caaaaac 1260
 Cp 4192 CAAAAAC 4133
 Db 1261 ttccgaagttaactgcttcaacacacacacacacacacacacacacacacacac 1320
 Cp 4132 TTCCGAAGGTAACCTGCTTCAACACACACACACACACACACACACACACACAC 4073
 Db 1321 ccttaattgaac 1380
 Cp 4072 GGTAGTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 4013
 Db 1381 tctctgtac 1440
 Cp 4012 TCTCTGTAC 3953
 Db 1441 gacataacttaac 1500
 Cp 3952 GACGATAGTACCGGATTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3893
 Db 1501 ccagcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1560
 Cp 3892 CCAGCTTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3833
 Db 1561 gcgccacgcttcccgagcagcagcagcagcagcagcagcagcagcagcagcagc 1620
 Cp 3832 GCGCGACGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3773
 Db 1621 caagac 1680
 Cp 3772 CAGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3713
 Db 1681 ggtttgcac 1740
 Cp 3712 GGTTTGCAATCT 3653
 Db 1741 tatgaaaaaac 1800
 Cp 3652 TATSSAAAAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3593
 Db 1801 ctacacatgcttct 1860

QY 5353 CAAATAGGGCTTCGGGCGACATTTCGGGAAAAAGTGGCAG 5392

Search completed: Tue Mar 16 03:14:30 1999
Job time : 1469 secs.

[W][O][R][L][D]
[D][I][S][T][R][I][B][U][T][I][O][N]
[R][I][G][H][T][S]
[B][Y]
[O][X][F][O][R][D]
[M][O][L][E][C][U][L][A][R]
[L][I][M][I][T][E][D]
[C][O][P][Y][R][I][G][H][T]
[1][9][9][3]

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPSrch_nn n.a. n a database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 23:40:59 1999; MasPar time 6848.69 Seconds
Tabular output not generated. 1409.884 Million cell updates/sec

Title: >US-09-020-716-6
Description: (15392) from US09020716 seq
Perfect Score: 5392
N.A. Sequence: 1 CTAATTTGTAAGCGTTAATA...ATTTCGCGAAAGTGCGAC 5392
Comp: GATTTAACATTCGCAATTAT...TAAGGGCGTTTTCACGGTG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895389244 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109

59gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_est30 27:gb_est31
28:gb_est32 29:gb_est33 30:gb_est34

Statistics: Mean 12.869; Variance 2.781; scale 4.627

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	564	12.3	711	26	AG004608 Homo sapiens genomic D	0.00e+00
2	564	12.3	711	27	AG010947 Homo sapiens genomic D	0.00e+00
3	557	12.2	721	27	AG011001 Homo sapiens genomic D	0.00e+00
4	557	12.2	721	25	AG004662 Homo sapiens genomic D	0.00e+00
5	535	11.8	718	26	AG004363 Homo sapiens genomic D	0.00e+00
6	535	11.8	718	27	AG010489 Homo sapiens genomic D	0.00e+00
7	530	11.7	698	26	AG003787 Homo sapiens genomic D	0.00e+00
8	530	11.7	698	27	AG009976 Homo sapiens genomic D	0.00e+00
9	524	11.6	696	27	AG009765 Homo sapiens genomic D	0.00e+00
10	524	11.6	696	26	AG003576 Homo sapiens genomic D	0.00e+00
11	524	11.6	720	28	AG013858 Homo sapiens genomic D	0.00e+00
12	524	11.6	720	25	AG009062 Homo sapiens genomic D	0.00e+00
13	619	11.5	747	26	AG007052 Homo sapiens genomic D	0.00e+00

C	14	609	11.3	644	28	AQ115013	CIT-HSP-2374C24.TF CIT	0.00e+00
C	15	603	11.2	692	27	AQ040352	CIT-HSP-237K21.TF CIT	0.00e+00
C	16	595	11.2	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	597	11.1	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	18	598	11.1	748	17	A1124281	1.53 Drosophila 8-12 h	0.00e+00
C	19	593	11.0	640	27	AQ074598	CIT-HSP-2382L21.TF CIT	0.00e+00
C	20	595	11.0	654	26	B78704	CIT78SK-A-880F6.TPC C	0.00e+00
C	21	594	11.0	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	22	588	10.9	652	28	AQ108633	CIT-HSP-2379M3.TF CIT	0.00e+00
C	23	583	10.8	594	26	C1786	Ciona intestinalis gen	0.00e+00
C	24	583	10.8	666	27	AQ079096	CIT-HSP-2356E24.TF CIT	0.00e+00
C	25	577	10.7	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.5	642	28	AQ113213	CIT-HSP-2372L22.TF CIT	0.00e+00
C	27	566	10.5	682	27	AQ074693	CIT-HSP-2381L23.TF CIT	0.00e+00
C	28	559	10.4	742	28	AQ108392	CIT-HSP-2383E2.TF CIT	0.00e+00
C	29	557	10.3	615	28	AQ112424	CIT-HSP-2376F2.TF CIT	0.00e+00
C	30	551	10.2	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.2	634	27	AQ039010	CIT-HSP-2356F2.TF CIT	0.00e+00
C	32	549	10.2	653	27	AQ077504	CIT-HSP-2365K21.TF CIT	0.00e+00
C	33	544	10.1	633	27	AQ040787	CIT-HSP-2335K22.TF CIT	0.00e+00
C	34	545	10.1	661	27	AQ057603	CIT-HSP-2345M2.TF CIT	0.00e+00
C	35	545	10.1	690	27	AQ029399	RPC111-39F1.TP RPC111	0.00e+00
C	36	540	10.0	637	27	AQ077360	CIT-HSP-2365E1.TF CIT	0.00e+00
C	37	534	9.9	594	26	B71605	RPC111-9N1.TP RPC111 H	0.00e+00
C	38	536	9.9	626	28	AQ111342	CIT-HSP-2371M2.TF CIT	0.00e+00
C	39	526	9.8	540	27	AQ074307	CIT-HSP-2382N1.TF CIT	0.00e+00
C	40	531	9.8	549	27	AQ093191	PF011-19F5.TP PF011	0.00e+00
C	41	526	9.8	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C	42	527	9.8	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C	43	530	9.8	630	27	AQ003238	RPC111-19F13.TP RPC111	0.00e+00
C	44	530	9.8	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C	45	525	9.7	600	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00

ALIGNMENTS

RESULT	1	AG004608	711 bp	DNA	GSS	29-JAN-1998
LOCUS		Homo sapiens genomic DNA, 21q region, clone: S594BG38				genomic
DEFINITION		survey sequence.				
ACCESSION		AG004608				
NID		92822058				
KEYWORDS		GSS.				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens DNA, clone: S594BG38.				
REFERENCE		1 (bases 1 to 711)				
AUTHORS		Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
TITLE		Homo sapiens genomic DNA, chromosome 21q				
JOURNAL		Published Only in Database (1998) In press				
REFERENCE		2 (bases 1 to 711)				
AUTHORS		Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.				
FEATURES		Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732, Fax: 0427-78-9561)				
BASE COUNT		169 a 188 c 172 g 182 t				
ORIGIN		Location/Qualifiers				
		1..711				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="21"				
		/clone="S594BG38"				
		/map="21q"				
		Query Match 12.3% Score 664; DB 26; Length 711;				
		Best Local Similarity 99.6%; Pred. No. 0.00e+00;				
		Matches 682; Conservative 0; Mismatches 0; Indels 3; Gaps 3;				

```

Db 24 TACATAATACGATACGAGAGAGTTTACCATCTGGCCGACGCTGCAATGATACAGCG 87
QY 4486 TACATAATACGATACGAGAGAGTTTACCATCTGGCCGACGCTGCAATGATACAGCG 4544
Db 88 AGACGACGCTGACGAGCTGACGATGATATATGATATATGATATATGATATATGATAT 147
QY 4545 AGACGACGCTGACGAGCTGACGATGATATATGATATATGATATATGATATATGATAT 4504
Db 148 GGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
QY 4605 GGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4564
Db 208 AGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
QY 4665 AGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4724
Db 268 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
QY 4725 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4784
Db 328 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387
QY 4785 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4844
Db 388 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 4845 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4904
Db 448 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 4905 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4964
Db 508 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 4965 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5024
Db 568 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 5085 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5143
Db 687 GTGACCCCAACTGATCTTCAGCATC 711
QY 5144 GTGACCCCAACTGATCTTCAGCATC 5168

RESULT 2 AG010947 711 bp DNA GSS 28-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: S594RG38, genomic
DEFINITION survey sequence.
ACCESSION AG010947
KEYWORDS g3347775
SOURCE GSS.
ORGANISM Homo sapiens
            Homo sapiens DNA, clone: S594RG38.
            Primates; Catarrhini; Hominidae; Homo.

REFERENCE
  1 (bases 1 to 711)
  AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
  TITLE Homo sapiens genomic DNA, chromosome 21q
  JOURNAL Published only in database (1999) In press
REFERENCE
  2 (bases 1 to 711)
  AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
  TITLE Direct Submission
  JOURNAL Submitted (28-Jun-1998) to the DDBJ/EMBL/GenBank databases
  Masahira Hattori, Kitasato University, Department of Science, JST
  Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
  (E-mail: hattori@shiga.kit-u.ac.jp, Tel: 0427-78-9732,
  Fax: 0427-78-9561)
  Location/Qualifiers
FEATURES

```

```

source 1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="S594RG38"
/map="21q"
BASE COUNT 159 a 189 c 172 g 182 t
ORIGIN
Query Match 12.4% Score 664; DB 27; Length 711;
Best Local Similarity 99.6% Pred. No. 0.00e+00;
Matches 682; Conservative 0; Mismatches 0; Indels 4; Gaps 0
Db 28 TAGATAAATACGATACGAGAGAGTTTACCATCTGGCCGACGCTGCAATGATACAGCG 27
QY 4486 TAGATAAATACGATACGAGAGAGTTTACCATCTGGCCGACGCTGCAATGATACAGCG 4544
Db 88 AGACGACGCTGACGAGCTGACGATGATATATGATATATGATATATGATATATGATAT 147
QY 4545 AGACGACGCTGACGAGCTGACGATGATATATGATATATGATATATGATATATGATAT 4504
Db 148 GGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
QY 4605 GGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4564
Db 208 AGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
QY 4665 AGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4724
Db 268 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
QY 4725 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4784
Db 328 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387
QY 4785 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4844
Db 388 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 4845 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4904
Db 448 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
QY 4905 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4964
Db 508 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 4965 TAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5024
Db 568 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 5085 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5143
Db 687 GTGACCCCAACTGATCTTCAGCATC 711
QY 5144 GTGACCCCAACTGATCTTCAGCATC 5168

RESULT 3 AG011001 721 bp DNA GSS 29-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: S79612X91, genomic
DEFINITION survey sequence.
ACCESSION AG011001
KEYWORDS g3357935
SOURCE GSS.
ORGANISM Homo sapiens
            Homo sapiens DNA, clone: S79612X91.
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.

```



```

REFERENCE
AUTHORS      1 (bases 1 to 721)
TITLE        Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL      Homo sapiens genomic DNA, chromosome 21q
PREFERENCE   Published Only in Database (1998) In press
AUTHORS      2 (bases 1 to 721)
TITLE        Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL      Direct Submission
SUBMITTED    Submitted (29-JUN-1998) to the DDBJ/EMBL/GenBank databases.
AUTHORS      Masahira Hattori, Kitasato University, Department of Science, JST
TITLE        Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
JOURNAL      (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
              Fax:0427-78-9561)
FEATURES
source       Location/Qualifiers
            1..721
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /clone="879G12X91"
            /map="21q"
BASE COUNT   195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 12.2%, Score 657, DB 27, Length 721.
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Lb 32 CATCTTACGGAAGGCAAGAGAAATATGCACTGCTGCCATTAACCATGAGTGAT 91
Cp 4937 CATCTTACGGAAGGCAAGAGAAATATGCACTGCTGCCATTAACCATGAGTGAT 4878
Db 92 AACACTGCGGCCAACTTACTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 151
Cp 4877 AACACTGCGGCCAACTTACTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 4818
Db 152 TTGCACAAATGGGGATCATGTAACTCGCTTGATGCTTGGGAACCGAGCTGAATGAA 211
Cp 4817 TTGCACAAATGGGGATCATGTAACTCGCTTGATGCTTGGGAACCGAGCTGAATGAA 4758
Db 212 GCCATACCAAGAGGAGCGGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCG 271
Cp 4757 GCCATACCAAGAGGAGCGGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCG 4698
Db 272 AAATCTTAATCTGGCAACTACTTACTTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 331
Cp 4897 AAATCTTAATCTGGCAACTACTTACTTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 4638
Db 332 GAGCGGATAAAGTTGAGCAACCACTTCTGCGCTGCGCCCTTCCGGCTGGCTGGTTTATT 391
Cp 4637 GAGCGGATAAAGTTGAGCAACCACTTCTGCGCTGCGCCCTTCCGGCTGGCTGGTTTATT 4578
Db 392 GCTGATAAATCTGGAGCCGTTGAGCTGAGTATCTACACGACGGGAGTCAGCAACTATGAT 451
Cp 4577 GCTGATAAATCTGGAGCCGTTGAGCTGAGTATCTACACGACGGGAGTCAGCAACTATGAT 4518
Db 452 GATGGTAAGCCCTCCCGTATCTAGTATCTACACGACGGGAGTCAGCAACTATGAT 511
Cp 4517 GATGGTAAGCCCTCCCGTATCTAGTATCTACACGACGGGAGTCAGCAACTATGAT 4458
Db 512 GAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAAGC-TTGGTAAGCTGCA 570
Cp 4457 GAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAAGC-TTGGTAAGCTGCA 4398
Db 571 GACCAAGTTTACTC-TATATACTTTAGATGATTTAAACTTCATTTTAAATTTAAAGG 629
Cp 4397 GACCAAGTTTACTC-TATATACTTTAGATGATTTAAACTTCATTTTAAATTTAAAGG 4338
Db 630 ATCTAGGTGAAGATCTTTT-GATAATCATGACCAAA-TCCCGCTAAGGTGAGTTTCG 687
Cp 4337 ATCTAGGTGAAGATCTTTT-GATAATCATGACCAAA-TCCCGCTAAGGTGAGTTTCG 4278
Db 688 TTCACCTGAGCGTCAGACCCCGTAGAAAAAT 719
Cp 4277 TTCACCTGAGCGTCAGACCCCGTAGAAAAAT 4246

```

```

4
RESULT       AG004662      721 bp      DNA      GSS      30-JAN-1998
LOCUS        Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic
DEFINITION   survey sequence.
ACCESSION    AG004662
NID          Q2826191
KEYWORDS     GSS.
SOURCE       Homo sapiens DNA, clone: 879G12X91.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria.
Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 721)
AUTHORS      Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE        Homo sapiens genomic DNA, chromosome 21q
JOURNAL      Published Only in Database (1998) In press
REFERENCE    2 (bases 1 to 721)
AUTHORS      Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE        Direct Submission
JOURNAL      Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.
AUTHORS      Masahira Hattori, Kitasato University, Department of Science, JST
TITLE        Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
JOURNAL      (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
              Fax:0427-78-9561)
FEATURES
source       Location/Qualifiers
            1..721
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /clone="879G12X91"
            /map="21q"
BASE COUNT   195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 12.2%, Score 657, DB 26, Length 721.
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGAAGGCAAGAGAAATATGCACTGCTGCCATTAACCATGAGTGAT 91
Cp 4937 CATCTTACGGAAGGCAAGAGAAATATGCACTGCTGCCATTAACCATGAGTGAT 4878
Db 92 AACACTGCGGCCAACTTACTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 151
Cp 4877 AACACTGCGGCCAACTTACTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 4818
Db 152 TTGCACAAATGGGGATCATGTAACTCGCTTGATGCTTGGGAACCGAGCTGAATGAA 211
Cp 4817 TTGCACAAATGGGGATCATGTAACTCGCTTGATGCTTGGGAACCGAGCTGAATGAA 4758
Db 212 GCCATACCAAGAGGAGCGGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCG 271
Cp 4757 GCCATACCAAGAGGAGCGGTGACACCAACGATGCTGTAGCAATGGCAACACGTTGGCG 4698
Db 272 AAATCTTAATCTGGCAACTACTTACTTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 331
Cp 4897 AAATCTTAATCTGGCAACTACTTACTTCTGACAAAGATCGGAGACCGAAGAGCTAACCGCTTTT 4638
Db 332 GAGCGGATAAAGTTGAGCAACCACTTCTGCGCTGCGCCCTTCCGGCTGGCTGGTTTATT 391
Cp 4637 GAGCGGATAAAGTTGAGCAACCACTTCTGCGCTGCGCCCTTCCGGCTGGCTGGTTTATT 4578
Db 392 GCTGATAAATCTGGAGCCGTTGAGCTGAGTATCTACACGACGGGAGTCAGCAACTATGAT 451
Cp 4577 GCTGATAAATCTGGAGCCGTTGAGCTGAGTATCTACACGACGGGAGTCAGCAACTATGAT 4518
Db 452 GATGGTAAGCCCTCCCGTATCTAGTATCTACACGACGGGAGTCAGCAACTATGAT 511
Cp 4517 GATGGTAAGCCCTCCCGTATCTAGTATCTACACGACGGGAGTCAGCAACTATGAT 4458
Db 512 GAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAAGC-TTGGTAAGCTGCA 570
Cp 4457 GAACGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAAGC-TTGGTAAGCTGCA 4398
Db 571 GACCAAGTTTACTC-TATATACTTTAGATGATTTAAACTTCATTTTAAATTTAAAGG 629
Cp 4397 GACCAAGTTTACTC-TATATACTTTAGATGATTTAAACTTCATTTTAAATTTAAAGG 4338
Db 630 ATCTAGGTGAAGATCTTTT-GATAATCATGACCAAA-TCCCGCTAAGGTGAGTTTCG 687
Cp 4337 ATCTAGGTGAAGATCTTTT-GATAATCATGACCAAA-TCCCGCTAAGGTGAGTTTCG 4278
Db 688 TTCACCTGAGCGTCAGACCCCGTAGAAAAAT 719
Cp 4277 TTCACCTGAGCGTCAGACCCCGTAGAAAAAT 4246

```



```

Db 141 CATACAAAATGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATAC 200
QY 3673 CATACAAAATGAGAGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATAC 3732
Db 201 CAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCCGCTTACC 260
QY 3733 CAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCCGCTTACC 3792
Db 261 GGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCGTGT 320
QY 3793 GGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCGTGT 3852
Db 321 AGGTATCTCAGTTCGGTGTAGTGTGCTTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCCC 380
QY 3853 AGGTATCTCAGTTCGGTGTAGTGTGCTTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCCC 3912
Db 381 GTTCAGCGGACCGCTGCGCCCTTATCCGGTAACCTATCGTTCGTGAGTCCAAACCCGTTAGA 440
QY 3913 GTTCAGCGGACCGCTGCGCCCTTATCCGGTAACCTATCGTTCGTGAGTCCAAACCCGTTAGA 3972
Db 441 CAGGACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCAGGTTATGTA 500
QY 3973 CAGGACTTATCGGCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCAGGTTATGTA 4032
Db 501 GGGGCTCTTAACAGAGTCTTTAAAGTGTGGCTTCAACTACGCTTAACACTAGAGAACAG 560
QY 4033 GGGGCTCTTAACAGAGTCTTTAAAGTGTGGCTTCAACTACGCTTAACACTAGAGAACAG 4090
Db 561 TATTGGTATCTCGCTGCTGCTGAAGCCAGTTTANGCTTTTCGGAAGAGAGTTGGTAGCTCT 620
QY 4091 TATTGGTATCTCGCTGCTGCTGAAGCCAGTTTANGCTTTTCGGAAGAGAGTTGGTAGCTCT 4149
Db 621 TGATCGGGCAACAAACCCAGCGCTGGTAGCG-IGTIIITITITITITGCAAGCAGCAGATA 679
QY 4150 TGATCGGGCAACAAACCCAGCGCTGGTAGCGTGGTITITITITGTTTGGCAAGCAGCAGATT 4209
Db 580 ACGGCGAGAAAAAAGAGTCTCAGA-GATC-TTTGATCTT 718
QY 4210 ACGGCGAGAAAAAAGAGTCTCAGAAGATCTTTGATCTT 4250

RESULT 7
LOCUS AG003787 698 bp DNA GSS 07-JAN-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
sequence.
ACCESSION AG003787
NID g2754689
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T485XN.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1998) to the DBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, tel:0427-78-9732,
Fax:0427-78-9561)
Location/Qualifiers
1. .698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="T485XN"
/map="21q"

BASE COUNT 189 a 153 c 171 g 175 t 10 others
ORIGIN

```

```

Query Match 11.7% Score 630: DB 26: Length 698;
Best Local Similarity 99.6% Pred No. n 00e+00:
Matches 661: Conservative 0: Mismatches 2: Indels 5: Gaps 5:

Db 31 CTTACGGATGGCATGACAGTAAGAGAATTATTCAGTCTGCCATTAACCATGAGTATAC 90
Cp 4934 CTTACGGATGGCATGACAGTAAGAGAATTATTCAGTCTGCCATTAACCATGAGTATAC 4875
Db 91 ACTGGGGCAACTTACTTCTGACACAGATCGAGAGCCGGAAGAGCTTAACCGCTTTTIS 150
Cp 4874 ACTGGGGCAACTTACTTCTGACACAGATCGAGAGCCGGAAGAGCTTAACCGCTTTTIS 4815
Db 151 CACAACATGGGGGATCATGTAACTCGCTTGTGTTGGGAACCGAGCTGAATGAAGCC 210
Cp 4814 CACAACATGGGGGATCATGTAACTCGCTTGTGTTGGGAACCGAGCTGAATGAAGCC 4755
Db 211 ATACCAAGCAGCGCGTGCACACAGATGCTGTAGCAATGGCAACACGTTCCGCAAA 270
Cp 4754 ATACCAAGCAGCGCGTGCACACAGATGCTGTAGCAATGGCAACACGTTCCGCAAA 4695
Db 271 CTATTAACCTGGGAACTACTTACTCTAGCTTCCGCGGAACAATTAATAGACTGATGAG 330
Cp 4694 CTATTAACCTGGGAACTACTTACTCTAGCTTCCGCGGAACAATTAATAGACTGATGAG 4635
Db 331 GCGGATAAAGTTGCAGGACCACTTCTCGGCTCGGCCCTTCGCGCTGGCTGGTTTATTGCT 390
Cp 4634 GCGGATAAAGTTGCAGGACCACTTCTCGGCTCGGCCCTTCGCGCTGGCTGGTTTATTGCT 4575
Db 391 GATAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATCTATTCGAGCACTGGGGCTAGAT 450
Cp 4574 GATAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATCTATTCGAGCACTGGGGCTAGAT 4515
Db 451 GGTAAAGCCCTCCGCTATCGTAGTATCTACACAGCGG-AGTCAGGCAACTATGATGAA 509
Cp 4514 GGTAAAGCCCTCCGCTATCGTAGTATCTACACAGCGGAGTCAAGCAACTATGATGAA 4455
Db 510 CGAAATAGACAGATTGCGTGAGATAGTGGCTCACTGTTAAGCATTGGTAACTGTAGA 569
Cp 4454 CGAAATAGACAGAT-CGCTGAGATAGTGGCTCACTGATTAAAGCATTGGTAACTGTAGA 4396
Db 570 CCAGTTTACTC-TATATACITTAGATTGATTAAAGCTTCATTTTAAATAAGGAT 628
Cp 4395 CCAGTTTACTC-TATATACITTAGATTGATTAAAGCTTCATTTTAAATAAGGAT 4336
Db 629 CTAGTGAAGATCCTTTT-GATAATCTCATGACCAATTCCTTAAGCTGAGTTTGGT 687
Cp 4335 CTAGTGAAGATCCTTTTGTATAATCTCATGACCAAAATCCCTTAAGCTGAGTTTGGT 4276
Db 688 TCCACTGA 695
Cp 4275 -CCACTGA 4269

RESULT 8
LOCUS AG009976 698 bp DNA GSS 06-JUL-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
sequence.
ACCESSION AG009976
NID g3289962
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T485XN.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) to the DBJ/EMBL/GenBank databases.

```


Cp 4607 GGTCTGAGGCTTT-CGGGCTG-CTAGTATATGATGATAAAATGAGAGGAGGCTG 4550
 Db 690 GGTCTC 695
 Cp 4549 GGTCTC 4544

RESULT 10
 LOCUS AG003576 696 bp DNA GSS 19-DEC-1997
 DEFINITION Homo sapiens genomic DNA, 21q region, clone. P8G4SpN15, genomic survey sequence.
 ACCESSION AG003576
 NID g2706702
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: P8G4SpN15.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 696)
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1997) In press
 REFERENCE 2 (bases 1 to 696)
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-1997) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)
 FEATURES
 source
 1. .696
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="P8G4SpN15"
 /map="21q"
 BASE COUNT 183 a 165 c 185 g 162 t 1 others
 ORIGIN
 Query Match 11 6% Score 624; DB 26; Length 696;
 Best Local Similarity 99.1%; Pred. No. 0.00e+00;
 Matches 660; Conservative 0; Mismatches 0; Indels 6; Gaps 6;

Db 31 GTTTTCTCACCAGGAGGCTGTGTAAGTAAAGATGCTGAAGATCAGTTGGTGCA 90
 Cp 5204 GTTTTCTCACCAGGAGGCTGTGTAAGTAAAGATGCTGAAGATCAGTTGGTGCA 5145
 Db 91 CGAGTGGTTACATCGAATCGATCTCAACAGCGGTAGATCCTTGAGAGTTTCGCCCC 150
 Cp 5144 CGAGTGGTTACATCGAATCGATCTCAACAGCGGTAGATCCTTGAGAGTTTCGCCCC 5085
 Db 151 GAAGAGGTTTCCATGATGAGCAGCTTTAAAGTTCTGCTATGTGGCGGCTATTATCC 210
 Cp 5084 GAAGAGGTTTCCATGATGAGCAGCTTTAAAGTTCTGCTATGTGGCGGCTATTATCC 5025
 Db 211 CGTATTGACCGGGCAAGAGCACTCGGTCCCGCATACACTATCTCAGATGACTTG 270
 Cp 5024 CGTATTGACCGGGCAAGAGCACTCGGTCCCGCATACACTATCTCAGATGACTTG 4965
 Db 271 GTTGAGTACTACAGCTACAGAAAAGATCTTACGGATGGCATCACAGTAAAGAAATTA 330
 Cp 4964 GTTGAGTACTACAGCTACAGAAAAGATCTTACGGATGGCATCACAGTAAAGAAATTA 4905
 Db 331 TGCAGTCTGCCATAACCATGAGTATACACTGCGGCCCACTTACTTCTGACACGATC 390
 Cp 4904 TGCAGTCTGCCATAACCATGAGTATACACTGCGGCCCACTTACTTCTGACACGATC 4845
 Db 391 GGAGAGCCGAAGAGCTATACCGTCTTCTGACACATGGGGATCATGTAACTCGCCTT 450
 Cp 4844 GGAGAGCCGAAGAGCTATACCGTCTTCTGACACATGGGGATCATGTAACTCGCCTT 4785
 Db 451 GATCGTTGGGAACCGGAGCTGAATGAAGGCATACCAACAGCAGCGGTGAACACCACT 510

Cp 4784 GATCGTTGGGAACCGGAGCTGAATGAAGGCATACCAACAGCAGCGTGA-CAACAGAI 4726
 Db 511 GCGCTGTAGCAATGGCAACAACCTTGGCAAACTATTAACTGGGCACTACTACTCTAG 570
 Cp 4725 G-CCTGTAGCAATGGCAACAACCTTGGCAAACTATTAACTGGGCACTACTACTCTAG 4567
 Db 571 CTTCGCGGGCAACAACTTAATAGACTGGATGGAGTGGGATAAAGTTTCAGGACCACTCTG 630
 Cp 4666 CTTCGCGG-CAACAATTAATAGACTGGATGGAGTGGGATAAAGTTTCAGGACCACTCTG 4608
 Db 631 CGCTCGGCGCTTTCGCGCTGGGTGTTTATTGCTGTATTAATCTGAGCGGCTGAGG-TC 689
 Cp 4607 CGCTCGGCGCTT-CGCGGTGG-CTGGTTTATTGCTGTATTAATCTGAGCGGCTGAGG-TC 4560
 Db 690 GGTCTC 695
 Cp 4549 GGTCTC 4544

RESULT 11
 LOCUS AG013858 720 bp DNA GSS 10-SEP-1998
 DEFINITION Homo sapiens genomic DNA, 21q region, clone. 762015SpN14, genomic survey sequence.
 ACCESSION AG013858
 NID g3560348
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: 762015SpN14.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 720)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 720)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)
 FEATURES
 source
 1. .720
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="762015SpN14"
 /map="21q"
 BASE COUNT 172 a 178 c 170 g 186 t 14 others
 ORIGIN
 Query Match 11 6% Score 624; DB 28; Length 720;
 Best Local Similarity 97.2%; Pred. No. 0.00e+00;
 Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGCTCCCTCGTCCGCTCCTCTGTCGACCTGCGGCTTACCGGATACCTGTCGCGCTT 94
 QY 3751 AGCTCCCTCGTCCGCTCCTCTGTCGACCTGCGGCTTACCGGATACCTGTCGCGCTT 3810
 Db 95 CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACCTGTAGTATCTCAGTTCGCTG 154
 QY 3811 CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACCTGTAGTATCTCAGTTCGCTG 3870
 Db 155 TAGTGTGTTTCGCTCAAGCTGGCGTGTGGCAGCAACCCCGCTTACGCGGCTGTC 214
 QY 3871 TAGTGTGTTTCGCTCAAGCTGGCGTGTGGCAGCAACCCCGCTTACGCGGCTGTC 3930
 Db 215 GCTTATCGGTAACATATGCTTTGAGTCCCAACCGGTAGACACGAGCTTATCGCCACTG 274
 QY 3931 GCTTATCGGTAACATATGCTTTGAGTCCCAACCGGTAGACACGAGCTTATCGCCACTG 3990


```

/chrnosome="21"
/clone="c48B2A23"
/map="21q"
BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 11.5% Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTGGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCGGC 100
QY ATCTGGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCGGC 4158
Db 101 AAACAACACAGCTGCTGAGCGTG-ITTTTGTGTTCAAGCAGCAGATTACCGGAGA 159
QY AAACAACACAGCTGCTGAGCGTGTTTGTGTTCAAGCAGCAGATTACCGGAGA 4218
Db 160 AAAAAGAGTCTCAAGAAGATCCTTTGATCTTTCTACGGGCTGAGCGCTCAGTGAAC 219
QY AAAAAGAGTCTCAAGAAGATCCTTTGATCTTTCTACGGGCTGAGCGCTCAGTGAAC 4278
Db 220 GAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 279
QY GAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 4338
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATATATAGTAAACTTGGTCT 339
QY CTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATATATAGTAAACTTGGTCT 4398
Db 340 GACAGTTACCAATGCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTCGTCA 399
QY GACAGTTACCAATGCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTCGTCA 4458
Db 420 TCAATAGTCTGACTCCCGTGGTGTATAGTATAGTATAGTATAGTATAGTATAGTAT 458
QY TCAATAGTCTGACTCCCGTGGTGTATAGTATAGTATAGTATAGTATAGTATAGTAT 4518
Db 459 GGGGCTAGTCTGTAATGATAGGAGAGGACGACGCTACCGGCTCCAGATTTATCAGCA 518
QY GGGGCTAGTCTGTAATGATAGGAGAGGACGACGCTACCGGCTCCAGATTTATCAGCA 4578
Db 519 ATAAACAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 577
QY ATAAACAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4637
Db 578 CATCCAGCTATTAAATGTTGCGGGAAGCTAGAGTAAAGTAGTTTCCGCGAGTTAATAGT 637
QY CATCCAGCTATTAAATGTTGCGGGAAGCTAGAGTAAAGTAGTTTCCGCGAGTTAATAGT 4695
Db 538 TTGGCAAGCTTTGTCATGCTGAGGAGGATGTCGTCACGCTCCGCTTCCGCTTAAAN 697
QY TTGGCAAGCTTTGTCATGCTGAGGAGGATGTCGTCACGCTCCGCTTCCGCTTAAAN 4754
Db 698 GGTTCATTTTCAGTCCGGTTCGCAACGATCAAGGGGTTTAAATGAT 746
QY GGTTCATTTTCAGTCCGGTTCGCAACGATCAAGGGGTTTAAATGAT 4801

RESULT 14
LOCUS A0115013 644 bp DNA GSS 28-AUG-1998
DEFINITION CIT-HSP-2374C24.TR CIT-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION A0115013
NID g3491134
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Adams, M. D., Rounsley, S. D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and

```

```

Venter, J. C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL
COMMENT
Unpublished (1998)

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/numcen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
Location/Qualifiers
source 1..644
/organism="Homo sapiens"
/clone="2374C24"
/db_xref="taxon:9606"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
BASE COUNT 135 a 183 c 145 g 179 t 2 others
ORIGIN

Query Match 11.3% Score 609; DB 28; Length 644;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCACGTGGCGCTGCTTTTACACGCTGCTGACTGGGAAAACCCCTGGCTTACCCACCTAAT 64
Cp 618 TCACGTGGCGCTGCTTTTACACGCTGCTGACTGGGAAAACCCCTGGCTTACCCACCTAAT 559
Db 65 CGCCTTGCAGCACATCCCGCTTTCCGCTAGCTGAGTAAAGAGGAGGAGGAGGAGGAG 124
Cp 558 CGCCTTGCAGCACATCCCGCTTTCCGCTAGCTGAGTAAAGAGGAGGAGGAGGAGGAG 459
Db 125 CGCCTTGCAGCACATCCCGCTTTCCGCTAGCTGAGTAAAGAGGAGGAGGAGGAGGAG 183
Cp 498 CGCCTTGCAGCACATCCCGCTTTCCGCTAGCTGAGTAAAGAGGAGGAGGAGGAGGAG 439
Db 184 TTAAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
Cp 438 TTAAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
Db 244 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Cp 378 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
Db 304 CAAGCTCTAAATCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
Cp 318 CAAGCTCTAAATCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
Db 364 CCCAAAAAATTTGATTAGGCTGATGTTACGCTAGTGGCCATCGCCCTGATAGACGCTT 423
Cp 258 CCCAAAAAATTTGATTAGGCTGATGTTACGCTAGTGGCCATCGCCCTGATAGACGCTT 199
Db 424 TTTCGCGCTTTGACGCTTGGAGTCCACGCTTTTAAATAGTGGACTCTTGTCCAAACTGGA 483
Cp 198 TTTCGCGCTTTGACGCTTGGAGTCCACGCTTTTAAATAGTGGACTCTTGTCCAAACTGGA 139
Db 484 ACAACTCAACCTATCTCGGCTCTATTCTTTGATTATAGGAGATTTTCCGATTTTCG 543
Cp 138 ACAACTCAACCTATCTCGGCTCTATTCTTTGATTATAGGAGATTTTCCGATTTTCG 79
Db 544 GCCTATTGTTAAAAAATGAGCTGATTTTAAACAAAATTTTAAACGGAATTTTAAACAAATA 603
Cp 78 GCCTATTGTTAAAAAATGAGCTGATTTTAAACAAAATTTTAAACGGAATTTTAAACAAATA 19

```

Db 604 TTAAGGTTAAATAATT 619
 Cp 18 TTAAGGTTAAATAATT 3

RESULT 15
 LOCUS AQ040352 692 bp. 1NA GSS 10-MAR-1998
 DEFINITION CIT-HSP-2327K21.TF CIT HSP Homo sapiens genomic clone 2327K21,
 genomic survey sequence.
 ACCESSION AQ040352
 NID 43306184
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominidae; Homo;
 1 (Bases 1 to 692)
 AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Sub.F., Wible,C., Shizuya,H.,
 Simon,M. and Venter,J.C.
 TITLE Use of a random RAP-End Sequence Database for Sequence-Ready Map
 Building
 JOURNAL Unpublished (1997)
 COMMENT Other GSSs: CIT-HSP-2327K21.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genome Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@igrr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tb/human/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.
 FEATURES
 source
 1..692
 Location/Qualifiers
 /organism="Homo sapiens"
 /note="vector: pBelorAC11; Sire_1: HindIII; Site_2:
 HindIII"
 /db_xref="taxon:9606"
 /clone="2327K21"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 BASE COUNT 158 a 193 c 186 g 155 t
 ORIGIN

Query Match 11.2% Score 603; DB 27; Length 692;
 Best Local Similarity 99.8%; Pred. No. 0,000+00;
 Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 88 GTTGGGTTAATGATGATAGTATTTTCTGTGTGAAATTTATGCGCTGACAAATTC 147
 QY 4227 GTTGGGTTAATGATGATAGTATTTTCTGTGTGAAATTTATGCGCTGACAAATTC 3286

Db 148 CACACAAACATAGCGCGGAGACATAAAGCTGAAGCTGGGTGGCTTAATGAGTGGCT 207
 QY 4287 CACACAAACATAGCGCGGAGACATAAAGCTGAAGCTGGGTGGCTTAATGAGTGGCT 3346

Db 208 AATTCACATTAATTTGCTTGTGATCATTTGGGCGGTCTTTCATCTGGGAAATCTGCTGGC 257
 QY 4347 AATTCACATTAATTTGCTTGTGATCATTTGGGCGGTCTTTCATCTGGGAAATCTGCTGGC 3406

Db 268 AGTTGATTAATGATGATAGTATTTTCTGTGTGAAATTTATGCGCTGACAAATTC 327
 QY 4407 AGTTGATTAATGATGATAGTATTTTCTGTGTGAAATTTATGCGCTGACAAATTC 3466

Db 328 GGGGTTCTTCTGATTAATTTGCTTGTGATCATTTGGGCGGTCTTTCATCTGGGAAATCTGCTGGC 387
 QY 4467 GGGGTTCTTCTGATTAATTTGCTTGTGATCATTTGGGCGGTCTTTCATCTGGGAAATCTGCTGGC 3526

Db 488 CTCACCTCAAAAGGCGGTAACTAAGGTTATGACACAAATTAAGGGATTAAGGGAGCAAAACA 447
 QY 3527 CTCACCTCAAAAGGCGGTAACTAAGGTTATGACACAAATTAAGGGATTAAGGGAGCAAAACA 4596

Db 448 TGTGAGCAAAAGGCGGACAAACAAAGGCGGCAAGAAAGGCAAGAAAGGCGGCTGCTTCT 509
 QY 3587 TGTGAGCAAAAGGCGGACAAACAAAGGCGGCAAGAAAGGCGGCTGCTTCTGCTTCT 6046

Db 508 TCCATAGGCTGGCGGCGGCGGCTGAGAGGATCAACAAATTCGAGAGTAAAGGCAAGGCTG 597
 QY 3647 TCCATAGGCTGGCGGCGGCGGCTGAGAGGATCAACAAATTCGAGAGTAAAGGCAAGGCTG 4706

Db 568 GAAACGCGGACAGCACTATATAACAAATTAAGGCTGCTTCTGCTTCTGCTTCTGCTTCT 627
 QY 3707 GAAACGCGGACAGCACTATATAACAAATTAAGGCTGCTTCTGCTTCTGCTTCTGCTTCT 4766

Db 628 CTCTGTGTTGCGGAGGCGGCTGAGAGGATCAACAAATTCGAGAGTAAAGGCAAGGCTG 687
 QY 4767 CTCTGTGTTGCGGAGGCGGCTGAGAGGATCAACAAATTCGAGAGTAAAGGCAAGGCTG 4829

Db 688 TGGCG 592
 QY 3827 TGGCG 3831

Search completed: Tue Mar 16 02:49:40 1999
 Job time : 11321 secs.

W P E R E H (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n a database search, using Smith-Waterman algorithm
Run on: Tue Mar 16 05:59:40 1999; MasPar time 7909 93 seconds
Tabular output not generated. 1545.492 Million cell updates/sec

Title: >US-09-020-716-7
Description: (1-5173) from US00020716 seq
Perfect Score: 5173
N A Sequence: 1 CTAATGTGAAGCGTTAATA ATTCCCGGAAAGTGGCAC 5173
Comp: GATTACATTCGCAATTAT TAAAGGGGCTTTCAGCGTG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 502357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb157

1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_lo 8:em_lo 9:em_lo 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vi

Database: genbank110

16:gb_bal 17:gb_bal 18:gb_htg 19:gb_in 20:gb_lo 21:gb_lo
22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl 26:gb_pl
27:gb_pr 28:gb_pr 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi

Statistics: Mean 12.610; Variance 7.779; scale 1.621

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	2228	43.1	2961	32	ARBL2KSP	PBluescript II KS(-) v 0.00e+00
2	2228	43.1	2961	32	ARBL2KSM	PBluescript II KS(-) v 0.00e+00
3	2228	43.1	2979	32	CVKSLIC	Ligation-independent c 0.00e+00
4	2228	43.1	4289	32	XXU35136	Plasmid pBSL197 cloning 0.00e+00
5	2228	43.1	9655	32	SYNPGR8V	Cloning vector pGR8, c 0.00e+00
6	2228	43.1	9750	32	AF025397	Expression vector pESP 0.00e+00
7	2228	43.1	9760	32	EVU86815	Expression vector pESP 0.00e+00
8	2228	43.1	10765	32	EVU87875	PESE-1 yeast expressio 0.00e+00
9	2224	43.0	2964	32	SYNBLKSPV	BlueScribe KS Plus clo 0.00e+00
10	2224	43.0	2964	32	SYNBLKSMV	BlueScribe KS Minus cl 0.00e+00
11	2224	43.0	2967	32	UC2449	Cloning vector pDIRECT 0.00e+00
12	2225	43.0	4412	32	EVCOPI12N	Expression vector pCOR 0.00e+00
13	2225	43.0	7474	22	A85522	Sequence 26 from Paten 0.00e+00

c	14	2224	43.0	8072	32	AF038566	Cloning vector pFUN, c 0.00e+00
	15	2221	42.9	2958	32	ARBLKSM	PBluescript KS(-) vect 0.00e+00
	16	2221	42.9	2958	32	ARBLKSP	PBluescript KS(-) vect 0.00e+00
	17	2213	42.8	4433	32	EVCOPI11N	Expression vector pCOR 0.00e+00
	18	2208	42.7	4267	32	PRS304	Yeast integrative vect 0.00e+00
	19	2208	42.7	4373	32	PRS306	Yeast integrative vect 0.00e+00
	20	2208	42.7	4443	32	PRS303	Yeast integrative vect 0.00e+00
	21	2208	42.7	4783	32	PRS314	Yeast integrative vect 0.00e+00
	22	2208	42.7	4887	32	PRS316	Yeast centromere vecto 0.00e+00
	23	2208	42.7	4967	32	PRS313	Yeast centromere vecto 0.00e+00
c	24	2208	42.7	5056	32	SCU64694	Yeast integrating vect 0.00e+00
	25	2208	42.7	5187	32	U34887	Yeast integrating vect 0.00e+00
c	26	2208	42.7	5504	32	PRS305	Yeast integrative vect 0.00e+00
	27	2208	42.7	5693	32	SCU64593	Yeast integrative vect 0.00e+00
	28	2208	42.7	6018	32	PRS315	Yeast centromere vecto 0.00e+00
c	29	2202	42.6	2976	32	CVU61229	Cloning vector pKRX, c 0.00e+00
	30	2202	42.6	3426	32	ASPGREEN1	Artificial sequences, 0.00e+00
	31	2205	42.6	3681	22	I15651	Sequence 6 from patent 0.00e+00
	32	2205	42.6	3681	22	A20702	PKSEL5 DNA sequence. 0.00e+00
	33	2202	42.6	3886	32	XXU35129	Plasmid pBSL141 clonin 0.00e+00
	34	2202	42.6	4051	32	XXU35126	Plasmid pBSL121 clonin 0.00e+00
	35	2202	42.6	4285	32	XXU35137	Plasmid pBSL199 clonin 0.00e+00
	36	2202	42.6	4480	32	XXU35134	Plasmid pBSL190 clonin 0.00e+00
	37	2202	42.6	4569	32	XXU35130	Cloning vector pBSL142 clonin 0.00e+00
	38	2202	42.6	4755	32	U93713	Cloning vector pBSL119 clonin 0.00e+00
	39	2202	42.6	4842	32	XXU35125	Plasmid pBSL119 clonin 0.00e+00
	40	2202	42.6	4893	32	U93714	Cloning vector pBSL121 clonin 0.00e+00
	41	2202	42.6	5035	32	XXU35128	Plasmid pBSL130 clonin 0.00e+00
	42	2205	42.6	5314	22	I15649	Sequence 4 from patent 0.00e+00
	43	2205	42.6	5314	22	A20700	PMTV1 DNA sequence. 0.00e+00
	44	2202	42.6	6042	32	U93718	Cloning vector pBSL12 0.00e+00
	45	2202	42.6	6868	32	U93719	Cloning vector pBSL122 0.00e+00

ALIGNMENTS

RESULT 1
LOCUS ARBL2KSP 2961 bp DNA circular SYN 10-MAY-1995
DEFINITION PBluescript II KS(+) vector DNA, phagemid excised from lambda ZAPII.
ACCESSION X52327
NID 958061
KEYWORDS artificial sequence; cloning vector; expression vector, vector, cloning vectors.
SOURCE Cloning vectors.
ORGANISM Artificial sequences; Cloning vectors.
REFERENCE 1 (bases 1 to 2961)
AUTHORS Thomas, E.A.
TITLE Direct Submission
JOURNAL
MEDLINE 88319944
REFERENCE 3 (bases 1 to 2961)
AUTHORS Altig-Rees, M.A. and Short, J.M.
TITLE PBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES Location/Qualifiers
source 1..2961
misc_feature 747 a 738 c 755 g 721 t
BASE COUNT 747 a 738 c 755 g 721 t
ORIGIN /note="phagemid pBluescriptII KS(+)"
Query Match 43.18; Score 2228; EB 22; Length 2961;
Best Local Similarity 100.00; P-Val No 0.00e+00;

QY 3006 GGGGTTGGGHTAAATCATGTTATAGTGTTCCTGTGTGAAATTTGTTATCCGCTCACAAAT 3065
Db TCCACAAACATACGAGCGGAGCATAAAGTGTAAAGCCTCGGCTGCTATAGTACAG 931
QY 3066 TCCACAAATACGAGCGGAGCATAAAGTGTAAAGCCTCGGCTGCTATAGTACAG 3125
Db CTAACCTCACATTAATTCGGTTGGCTCACTGCGCGCTTCCAGTCGGGAAACCTGCTGTG 991
QY 3126 CTAACCTCACATTAATTCGGTTGGCTCACTGCGCGCTTCCAGTCGGGAAACCTGCTGTG 3185
Db CGAGCTGCATTAATGATCGGCAAGCGCGGAGAGCGGTTTCGCTATTCGGCGCTC 1051
QY 3186 CGAGCTGCATTAATGATCGGCAAGCGCGGAGAGCGGTTTCGCTATTCGGCGCTC 3245
Db TTCGCTTCCTCGCTCACTGAGCTCGCTCGGCTCGGCTTCGCTGCGGAGCGGTATC 1111
QY 3246 TTCGCTTCCTCGCTCACTGAGCTCGCTCGGCTTCGCTGCGGAGCGGTATC 3305
Db AGCTCACTCAAGCGGCTTAATAGCGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAA 1171
QY 3306 AGCTCACTCAAGCGGCTTAATAGCGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAA 3365
Db CATGTGAGCAAAAGCGCGCAAGGCGCAGGACCGTAAAGAGCGCGTTCGCTGCGCTT 1231
QY 3366 CATGTGAGCAAAAGCGCGCAAGGCGCAGGACCGTAAAGAGCGCGTTCGCTGCGCTT 3425
Db TTCCATAGGCTCGGCGCCCTCAAGGAGCATCACAAAATCAGAGCTCAAGTCAGAGGTG 1291
QY 3426 TTCCATAGGCTCGGCGCCCTCAAGGAGCATCACAAAATCAGAGCTCAAGTCAGAGGTG 3485
Db GCGAAACCCGACAGGACTATTAAGATACGAGGCGTTTCGCCCTGGAAGCTCCCTCGTGGC 1351
QY 3486 GCGAAACCCGACAGGACTATTAAGATACGAGGCGTTTCGCCCTGGAAGCTCCCTCGTGGC 3545
Db CTCCTCTGTTCCGAGCTCGCGCTTACCGGATACCTTCGCGCTTTCCTCTCGGGAAG 1411
QY 3546 CTCCTCTGTTCCGAGCTCGCGCTTACCGGATACCTTCGCGCTTTCCTCTCGGGAAG 3605
Db CGTGGCGCTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGTGTGCTTCGCTC 1471
QY 1472 CAAGCTGGGCTGTGTCACGAACCCCGCTTACGCCGACCGCTGGCGCTTATCCGGTAA 1531
QY 3666 CAAGCTGGGCTGTGTCACGAACCCCGCTTACGCCGACCGCTGGCGCTTATCCGGTAA 3725
Db CTATCGCTTTCAGTCCAAACCCGGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGG 1591
QY 3726 CTATCGCTTTCAGTCCAAACCCGGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGG 3785
Db TAACAGGATTAGCAGAGCGAGGTATAGCGCGGTGCTACAGAGTCTTGAAGTGGTGCC 1651
QY 3786 TAACAGGATTAGCAGAGCGAGGTATAGCGCGGTGCTACAGAGTCTTGAAGTGGTGCC 3845
Db TAACAGGATTAGCAGAGCGAGGTATAGCGCGGTGCTACAGAGTCTTGAAGTGGTGCC 1711
QY 3846 TAACAGGATTAGCAGAGCGAGGTATAGCGCGGTGCTACAGAGTCTTGAAGTGGTGCC 3905
Db CTTCGGAAGAGAGTTGGTAGCTCTTATCGCGCAACCAACACCGCTGGTAGCGGTGG 1771
QY 3906 CTTCGGAAGAGAGTTGGTAGCTCTTATCGCGCAACCAACACCGCTGGTAGCGGTGG 3965
Db TTTTTCCTGTTTGAAGCAGCAGATTACGCGCAGAAAGAGATCTCAAGAGATCCCTTT 1831
QY 3966 TTTTTCCTGTTTGAAGCAGCAGATTACGCGCAGAAAGAGATCTCAAGAGATCCCTTT 4025
Db GATCTTTCAGCGGCTCTCAGCTCAGTGGAGCAAGAACTCACGTTAAGGATTTTGGT 1891
QY 4026 GATCTTTCAGCGGCTCTCAGCTCAGTGGAGCAAGAACTCACGTTAAGGATTTTGGT 4085
Db CATGAGATTATCAAAAAGGATCTTCACTAGATCCCTTTTAAATTAAGATTTTAA 1951
QY 4086 CATGAGATTATCAAAAAGGATCTTCACTAGATCCCTTTTAAATTAAGATTTTAA 4145

Db 1952 ATCAATCTAAAGTATATATGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTGA 2011
QY 4146 ATCAATCTAAAGTATATATGATTAACCTTGGTCTGACAGTTACCAATGCTTAAATCAGTGA 4205
Db GGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCCATAGTTGCTGACCTCCCGCTCGT 2071
QY 4206 GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCCATAGTTGCTGACCTCCCGCTCGT 4265
Db GTAGATAACTAGATACGAGGAGGCTTACCAATCTGCGCTCAGTGTGCAATGATACGAGG 2131
QY 4266 GTAGATAACTAGATACGAGGAGGCTTACCAATCTGCGCTCAGTGTGCAATGATACGAGG 4325
Db AGACCCACGCTCAGCGGCTCGAGATTTATCAACAATAAAGCAGCAGTGGGAAGGAGGGA 2191
QY 4326 AGACCCACGCTCAGCGGCTCGAGATTTATCAACAATAAAGCAGCAGTGGGAAGGAGGGA 4385
Db GGCAGAGGTGCTGCAACTTTATCGCGCTTCCATCCAGTCTATTAATTTGCTGCCGGA 2251
QY 4386 GGCAGAGGTGCTGCAACTTTATCGCGCTTCCATCCAGTCTATTAATTTGCTGCCGGA 4445
Db AGCTCAGTAACTGCTGCGGAGTTAAAGTTTGGGCAACGTTTGTSCATTTGTAAGS 2311
QY 4446 AGCTCAGTAACTGCTGCGGAGTTAAAGTTTGGGCAACGTTTGTSCATTTGTAAGS 4505
Db CATGCTGCTCAGCGCTCGCTTGGTATGGCTTCACTGAGCTCCGCTTCCCAACGATC 2371
QY 4506 CATGCTGCTCAGCGCTCGCTTGGTATGGCTTCACTGAGCTCCGCTTCCCAACGATC 4565
Db AAGCGAGTTACATGATGCGGATGTTGCAAAAAGGCTTGGCTTCTGCGCTGCTGCTGCT 2431
QY 4566 AAGCGAGTTACATGATGCGGATGTTGCAAAAAGGCTTGGCTTCTGCGCTGCTGCTGCT 4625
Db GATCGTTGTCAGAACTAAATGCTGCGGAGTGTATCATCATGATGTTATGCGAGCAGTGA 2491
QY 4626 GATCGTTGTCAGAACTAAATGCTGCGGAGTGTATCATCATGATGTTATGCGAGCAGTGA 4685
Db TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2551
QY 4686 TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4745
Db CAAGTCATCTGAGAACTAGTATGCGGCGAGCGAGTTTGGCTTCTGCGCTGCTGCTGCTGCT 2611
QY 4746 CAAGTCATCTGAGAACTAGTATGCGGCGAGCGAGTTTGGCTTCTGCGCTGCTGCTGCTGCT 4805
Db GGATAATACCGGCGCAGATAGCAGAACTTTAAAGTGTCTCATCATGATGTTATGCGAGTGA 2671
QY 4806 GGATAATACCGGCGCAGATAGCAGAACTTTAAAGTGTCTCATCATGATGTTATGCGAGTGA 4855
Db GGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 2731
QY 4866 GGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 4925
Db TGCACCGAACTGATCTTACGATCTTTTACCTTCCAGCGCTTCTGCGTGGAGGAGGAGG 2791
QY 4926 TGCACCGAACTGATCTTACGATCTTTTACCTTCCAGCGCTTCTGCGTGGAGGAGGAGG 4985
Db AGGAAGCGAAATGCGGCAAAAAGGAAATAGGCGCAGCGGAAATGTGAAATCTCAT 2851
QY 4986 AGGAAGCGAAATGCGGCAAAAAGGAAATAGGCGCAGCGGAAATGTGAAATCTCAT 5045
Db ACTCTTCCCTTTTCAATATTTTGAAGCATTTATCAGGCTTATGCTCATGAGCGGATA 2911
QY 5046 ACTCTTCCCTTTTCAATATTTTGAAGCATTTATCAGGCTTATGCTCATGAGCGGATA 5105
Db CATATTTGAATGATTTAGAAAATTAACAAATAGGCTTCCGCGCAGATTTCCCGGAA 2971
QY 5106 CATATTTGAATGATTTAGAAAATTAACAAATAGGCTTCCGCGCAGATTTCCCGGAA 5165
Db AGTGCCAC 2979
QY 5166 AGTGCCAC 5173

```

RESULT      4      XX045136      4289 bp      DNA      circular      SYN      26-SEP-1995
LOCUS      plasmid pBSL97 cloning vector, complete sequence.
ACCESSION  U36136
VERSION    1
KEYWORDS   cloning vector pBSL97.
SOURCE     Cloning vector pBSL97.
ORGANISM   artificial sequence: cloning vectors.
REFERENCE  1 (bases 1 to 4289)
AUTHORS   Alexeyev,M.F., Shokolenko,I.N. and Crouchan,T.P.
TITLE     Improved antibiotic resistance gene cassette and mpaA elements
          for Escherichia coli vector construction and in vitro
          deletion/insertion mutagenesis
          Gene 160 (1): 63-67 (1995)
JOURNAL    9554958
MEDLINE    9554958
REFERENCE  2 (bases 1 to 4289)
AUTHORS   Henden,P.N.
TITLE     Direct Submission
JOURNAL    Submitted (31-Aug-1995) Paul N. Henden, Laboratory of Mathematical
          Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES   Location,Qualifiers
           1..4289
           /organism="Cloning vector pBSL97"
           /plasmid="pBSL97"
           /db_xref="taxon:42709"
           complement(804..1598)
           /codon_start=1
           /transl_table=11
           /product="neomycin phosphotransferase"
           /db_xref="PID:g984924"
           /translation="MIPFCILHATSSAAVKEFLFVFWA-JTLIGSSUAAVFLSAUGR
           PVLEVKTDLSALNLODEAARLSWLATGVPCCAALVDVYTEAGRDWLLGVEPQDIL
           LSSHIAAEKYSIMADARRRLHIDFATCFPDHQAARHTEFAPTEMEAGLSVVRULLE
           HQGLAPAFIPARIKAMPDPCFIVTHGDACIPNIMVENGSPFPGIDGRLGVADRY
           QDIALATPDIAEELGGSWADPFLVLYGIAAPDSQRIAFYELDEFF"
           complement(3301..4161)
           /EC_number="3.5.2.6"
           /codon_start=1
           /transl_table=11
           /function="ampicillin resistance"
           /product="beta-lactamase"
           /db_xref="PID:g984925"
           /translation="MSTQHPFVALIPFFAFLPVEFAHPETILVKVDARNDLQARVGY
           IELDLSKGLIESEPPPEPFPMMSTKVLICGAVLSPIIDAGOPOLGPRTHYSQNDLVE
           YSPVTEKHLTGMTVREPLQSAATIMSLETAANLLELIIIGSKELTAFLNHWKHVTRL
           DRWEPLENAIPNBERDTTPVAMATTLKRLTGELLTLASRQQLLHREADKVGPL
           LRSALPAPWFTANKSGAGFEMGSPRIITAAIGPPKPSPIVVITTSQATMPPNPQIA
           EIGASLKHW"
BASE COUNT 1024 a 1157 c 1111 g 947 t
ORIGIN
Query Match 43.1%; Score 2228; DB 32; Length 4289;
Best Local Similarity 100.0%; Prod No 8 006+00;
Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2062 ATGACATTTAGAGGGGGGCTGGTAATCAATTTTCTGCTGCTTATATGAGGGTTAAATTC 2121
QY 2446 GTCCACATCCAGGGGGGGGGGGTACCCAGCTTTTGCTCCCTTTAGTCAGGGTTAATTC 3005

DB 2122 GATATTTGGGTAAATATGATATATAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2181
QY 3006 GGGGTGGGTAAATATGATATATAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3065

DB 2182 TCCACACATACACAGCGGGAAGATATAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAG 2241
QY 3066 TCCACACATACACAGCGGGAAGATATAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAG 3125

DB 2242 CTAACATCATTAATTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2301
QY 3126 CTAACATCATTAATTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3185

```

```

DB 2302 CCAAGTTTCATTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2461
QY 3196 CCAAGTTTCATTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2465
DB 2362 TTCGGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2421
QY 3246 TTCGGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2405
DB 2422 AGTCATCTCAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2481
QY 3306 AGTCATCTCAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2465
DB 2492 CAGCTGACCAAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2541
QY 3366 CAGCTGACCAAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2425
DB 2542 TTTCATAGGCTTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2601
QY 3426 TTTCATAGGCTTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2485
DB 2602 CGCAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2661
QY 3486 CGCAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2465
DB 2662 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2721
QY 3546 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2605
DB 2722 CAGCTGACCAAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2781
QY 3606 CAGCTGACCAAAAAGGGGGGTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2665
DB 2782 CAAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2841
QY 3686 CAAAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2725
DB 2842 CTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2901
QY 3726 CTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2785
DB 2902 TAACAGGATTAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2961
QY 3786 TAACAGGATTAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2845
DB 2962 TAATTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3021
QY 3846 TAATTAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2905
DB 3022 GTTTCGAAAAAAGAGTTTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3081
QY 3906 GTTTCGAAAAAAGAGTTTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2965
DB 3082 TTTTITTTTITTTGAAAGTAGGATTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3141
QY 3966 TTTTITTTTITTTGAAAGTAGGATTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3025
DB 3142 CATCTTTTCTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3201
QY 4026 CATCTTTTCTTAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3085
DB 3202 CATGAGATTATCAAAAAGGAGCTTTCACCTACATGCTCTTTTAAATTAATAAATGAAGTTTAA 3261
QY 4086 CATGAGATTATCAAAAAGGAGCTTTCACCTACATGCTCTTTTAAATTAATAAATGAAGTTTAA 3145
DB 3262 ATCAATCTAAGATATATATGAGTAAATCTGCTGACAGCTTTCACCTACATGCTCTTAAATGAAG 3321
QY 4146 ATCAATCTAAGATATATATGAGTAAATCTGCTGACAGCTTTCACCTACATGCTCTTAAATGAAG 3205
DB 3322 GGCACCTATCTCAGGAGCTCTGCTTATTTGCTTCATCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT 3381
QY 4206 GGCACCTATCTCAGGAGCTCTGCTTATTTGCTTCATCATAGTTGCTGCTGCTGCTGCTGCTGCTGCT 3265
DB 3382 GTACATTAACATGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3441

```

```
|||||
QY 4266 GTAGTAACCTACGATACGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG 4325
Db 3442 AGACCCACGCTCACCGGCTCCAGATTATACAGCAATAAACACGAGCGCGGAAGGGCGGA 3501
QY 4326 AGACCCACGCTCACCGGCTCCAGATTATACAGCAATAAACACGAGCGCGGAAGGGCGGA 4385
Db 3502 GCGCAGAAGTGGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGGA 3561
QY 4386 GCGCAGAAGTGGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGGA 4445
Db 3562 AGCTAGAGTAAGTGGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGGA 3621
QY 4446 AGCTAGAGTAAGTGGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGGA 4505
Db 3622 CATCGTGGTGTACGCTCGCTGTTGGTATGGCTTCATTCAGCTCCGCTCCACAGCATC 3681
QY 4506 CATCGTGGTGTACGCTCGCTGTTGGTATGGCTTCATTCAGCTCCGCTCCACAGCATC 4565
Db 3682 AAGCGGAGTTACATGATCCGCCCTGTTGTGCAAAAAGCGGTAGCTCCCTCGGTCCTCC 3741
QY 4566 AAGCGGAGTTACATGATCCGCCCTGTTGTGCAAAAAGCGGTAGCTCCCTCGGTCCTCC 4625
Db 3742 GATCGTGTGCAAGTAAGTGGCCGCACTGTTATCACTCATCTGTTATGGCAGCACTGCA 3801
QY 4626 GATCGTGTGCAAGTAAGTGGCCGCACTGTTATCACTCATCTGTTATGGCAGCACTGCA 4685
Db 3802 TAATCTCTTACTGTCATGCACTCGCTAGAGTCTTTCTGTCGACTGGTGTAGTACTCAAC 3861
QY 4686 TAATCTCTTACTGTCATGCACTCGCTAGAGTCTTTCTGTCGACTGGTGTAGTACTCAAC 4745
Db 3862 CAAGTCATCTGAGAATAGTGTATGCGGCGACGAGTGTCTCTTCCCGGGCTCAATACG 3921
QY 4746 CAAGTCATCTGAGAATAGTGTATGCGGCGACGAGTGTCTCTTCCCGGGCTCAATACG 4805
Db 3922 GGATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGTTGAAACAGTCTTTC 3981
QY 4806 GGATATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCTGTTGAAACAGTCTTTC 4865
Db 3982 GGGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCG 4041
QY 4866 GGGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATTAACCCACTCG 4925
Db 4042 TGCACCACTGATCTTACGATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAAAC 4101
QY 4926 TGCACCACTGATCTTACGATCTTTTACTTTCACGAGCTTTCTGGGTGAGCAAAAAC 4985
Db 4102 AGGAAGCAAAATGCCGCAAAAAGGGAATAAGGCGGACACGGAATGTTGAATCTCAT 4161
QY 4986 AGGAAGCAAAATGCCGCAAAAAGGGAATAAGGCGGACACGGAATGTTGAATCTCAT 5045
Db 4162 ACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGCTTTATGCTCATGAGCGGATA 4221
QY 5046 ACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGCTTTATGCTCATGAGCGGATA 5105
Db 4222 CATATTGAATGATTATTAGAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA 4281
QY 5106 CATATTGAATGATTATTAGAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA 5165
Db 4282 AGTGCCAC 4289
QY 5166 AGTGCCAC 5173
```

```
RESULT 5
LOCUS SYNQR8V 9655 bp DNA circular SYN 04-SEP-1997
DEFINITION Cloning vector pGR8, complete sequence.
ACCESSION AB001531
NID g2340041
KEYWORDS
SOURCE Cloning vector pGR8.
ORGANISM Cloning vector pGR8.
          artificial sequence; cloning vectors.
```

```
REFERENCE 1 (bases 1 to 9655)
AUTHORS Mazzarella R. and Pillia, G.
TITLE Recombination trapping: an 'in vivo' approach to recover cDNAs
        encoded in YACs
JOURNAL Unpublished (1997)
COMMENT GSDB:S:127444.
FEATURES
    Location/Qualifiers
        1..9655
            /organism="Cloning vector pGR8"
            /plasmid="Cloning vector pGR8"
            /db_xref="taxon:64712"
        misc_feature 1..135
            /standard_name="telomere"
        prim_transcript 145..2097
            /gene="HIS5"
            /db_xref="PIR:g2358009"
        gene 145..2097
            /gene="HIS5"
        CDS 761..1915
            /gene="HIS5"
            /EC_number="2.6.1.9"
            /codon_start=1
            /transl_table=11
            /product="histidinol phosphate aminotransferase"
            /translation="MVFDLKRIVPRKIYNI:EPYFCAPDDFTEGIIIDANENAHSPVPV
            ELSKINLHPYDPHQLEFKTAMKYPNKTSYANDPEVKPLTADNLCIAGVSGDES:DA
            IIRACCVREKILVLPPTYSMSSVCANINDIEVQCPLIVSGSFOMDTEAVITILKN
            DSLIKLMEVTPSGNPTGAKIKTSLIEKVLONWNGLVVDEAYVDFGSGSTALVTKY
            PNLVLTQLSKSGFAGIRLGMTATYAEALRIINAMKAPINISLSEALYALKAVODSN
            LKMEATSKIIINEKMLLKTALDYVDQYVGGDLNELLIRINGNGNVLAKK:YY
            QLATQSGVYVFPFGNLCGSCGIPITVGTTFEENTHLIKYFKETLYLANE"
            /transcript 2138..6708
            /gene="LYS2"
            /db_xref="PIR:g2358010"
            /transl_table=11
            /product="alpha-aminoadipate reductase"
            /db_xref="PIR:g2358010"
            /translation="MTNEKVMIEKLDNPTLSVLPHDFLRPOQEPYTKOATYSLQLPOL
            DVPDFSNKYAVALSVAAL:YRVTGGDDIVLYIANNKILRFNIQPTWSENFELYSII
            NNELNKINSIEANFSDELAEKIOSCODLERPOLFLAFLENDCKLDFKHLVDF
            ALNDTSNNAHVLNLIYNSLLYSNERTIVADOFYVLAALSDPNCITKSLITAS
            SKDSLPOPTKLGKDFVGCIDHIFODNAEFPEKTCVETPTILNSKSPSYRDLIN
            RTSNIVAHYLIKTGKRGDVMYIYSSRGVDMVCMGVLRAGATFVIDPAYPAROI
            IYLGVAHRLIVIRAAQDLQVDEYINDELEIVSRINISIAIOENSTIEGKLNKE
            DVLAPYDRYKDTRTGVVGPDSNPTLSFTSGSEIPKGVLRHFLAYFNWMSKRFN
            LTENDKFTMLSGIAHDPIQDMFTPLFGLAQLYVPTODDITGTPRLAEMSKYGTVT
            HLTPAMGOLLTAQATTIPKLRHAFVGDILTKPCLRIQTLAENP:IVNVY:ITTEIQ
            RAVSYFEVKSNDPDLNKLKLVMPAGKMLNVLVNVNDPTOIGGIGEYLV
            RAGLAEGYRGLPELNKEKFNWVEKDHNYLDKNDGEPWRFWGLPDRILYRAGD
            LRTLPNPFDDKDLQSKQSVKEVETPIVKGLISQHLISKDITF:KPI:ASTAM
            LITVMPFPDKPDLQSKQSVKEVETPIVKGLISQHLISKDITF:KPI:ASTAM
            PSILVMDKLPNPKGDKPKQFTPKQNLVAENTVSETDSDFTNVERSVRLK
            LSILPTKPSVSDPSDFDLGHSILATKMI:FTLKKLOVDLPLGTFIKYPT:KAA:AA
            BIDRTKSSGSSGGEVENVNTANYAEDAKKIVETLPSSYPSREVFPEPNAEGKT:IN
            VFTGVTGFLGSY:LAOLLGSPKANSYKVFVAHPAKDEEAAPLOKAGITGTVNE
            KFSNKKVVLGDLKSGQFGLSDERKMDLANITVDIIHNGALVHVHVPYAKLRPNVIS
            TINYMSLAAGKPKFFDFVSTSTLDEYIFNLSKLVSEKPGILSEDC:MSASGL
            TGGYQSKAAEYIIRPAGEPGLPGCIVRPYVYVTSAGANGSNTDOLFLPLKGSVQL
            GKIPDIENSVMNPVDHVARVATSLNPKNELAVALAQTGHPRLIKDYLVLHUY
            GYDVEISYKWKKSLEASVIDRNEALYPLLHMLVDNLPESTKAPELDRAVASL
            KKDRTAWTGDMSNGIGVTPVEVGIYIAFLNKNVGLPPTTHNDKLPLPSIELTQAQ:SL
            VASGAGAGSSAAA"
            complement(6709..7185)
            /standard_name="f(+) origin"
        gene 7186..7496
            /gene="LacZ"
        prim_transcript 7186..7496
            /gene="LacZ"
```

[illegible]


```

QY 4866 TAATTCCTTACTGTCATGCAATCCGTAAGATGCTTTCTGTGATGGTGTAGTACTCAAC 4745
D 9220 CAAGTCATCTGAGATAGTGTATCGGCGACGAGTGTCTCTTCCCGGGGTCAATACG 9279
QY 4746 CAAGTCATCTGAGATAGTGTATCGGCGACGAGTGTCTCTTCCCGGGGTCAATACG 4805
D 9280 GGATAATACGCGCCACATAGCAGAACTTTAAAGTGTCTATCTTGGAAACGTTCTTC 9339
QY 4806 GGATAATACGCGCCACATAGCAGAACTTTAAAGTGTCTATCTTGGAAACGTTCTTC 4865
D 9340 GGGGCGCAAACTCTCAGGATCTTACGCTGTGTGATGATCCAGTGTGATCCCACTCG 9399
QY 4866 GGGGCGCAAACTCTCAGGATCTTACGCTGTGTGATGATCCAGTGTGATCCCACTCG 4925
D 9400 TGCACCACTGATCTTACGATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAC 9459
QY 4926 TGCACCACTGATCTTACGATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAC 4985
D 9450 AGGAAGCAAAATGCCGCAAAAGGAATAGGGCGGACCGAAATGTTGAATACTCAT 9519
QY 4986 AGGAAGCAAAATGCCGCAAAAGGAATAGGGCGGACCGAAATGTTGAATACTCAT 5045
D 9520 ACTCTTCCTTTTCAATATTAATTAAGCATTTATCAAGGTTATGTCATCAGCGGATA 9579
QY 5046 ACTCTTCCTTTTCAATATTAATTAAGCATTTATCAAGGTTATGTCATCAGCGGATA 5105
D 9580 CATATTTGAATGTTATAGAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGGAAA 9639
QY 5106 CATATTTGAATGTTATAGAAAATTAACAAATAGGGGTTCCGCGCACATTTCCCGGAAA 5165
D 9640 AGTCCAC 9647
QY 5166 AGTCCAC 5173

RESULT 5 AF025397 9750 bp DNA SYN 17-DEC-1997
LOCUS Expression vector pESP-3, complete sequence.
DEFINITION AF025397
ACCESSION AF025397
NID q2689634
KEYWORDS Expression vector pESP-3.
SOURCE Expression vector pESP-3.
ORGANISM Expression vector pESP-3.
REFERENCE 1 (bases 1 to 9750)
AUTHORS Lu,Q.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1997) Marketing Analysis, Stratagene, 11011 North
Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT On Dec 17, 1997 this sequence version replaced gi:2623066.
FEATURES
    source
        1..9750
            /organism="Expression vector pESP-3"
            /specific_host="Schizosaccharomyces pombe"
            /db_xref="taxon:67775"
            /plasmid="pESP-3"

BASE COUNT 2919 a 1959 c 1983 g 2889 t
ORIGIN
Query Match 43.1%; Score 2228; DB 32; Length 9750;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D 1273 GTGSCATTTTCGGGAAATGTGGGGAAACCCCTATTGTTTATTTTCTAAATACATT 1332
Cp 5173 GTGGCACTTTTCGGGAAATGTGGGGAAACCCCTATTGTTTATTTTCTAAATACATT 5114
D 1333 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTAAGAAA 1392
Cp 5113 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTAAGAAA 5054
D 1393 GGAAGAGTATGATTAATCAACATTTCCGTGTGCGCCCTATTCCCTTTTGTGGCGCATTT 1452

```

```

Cp 5053 GGAAGAGTATGATTAATCAACATTTCCGTGTGCGCCCTATTGCTTTTGTGGTATTTT 4964
D 1453 GCTTCCTCTTTTTCCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGCTGAACATCAGT 1512
Cp 4993 GCTTCCTCTTTTTCCTCACCAGAAACGCTGGTGAAGTAAAGATGCTGCTGAACATCAGT 4934
D 1513 TGGGTCCAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATGCTTCAAGATT 1572
Cp 4933 TGGGTCCAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATGCTTCAAGATT 4874
D 1573 TTGCCCCCAAGAAACGTTTTCATGATGAGCACTTTTAAATTTCTGCTATGCTGGCGGG 1632
Cp 4873 TTGCCCCCAAGAAACGTTTTCATGATGAGCACTTTTAAATTTCTGCTATGCTGGCGGG 4814
D 1633 TATTATCCCGTATTGACGGCGGCAAGCAACTCGTGGCGGCATACATATTTCTGAGA 1692
Cp 4813 TATTATCCCGTATTGACGGCGGCAAGCAACTCGTGGCGGCATACATATTTCTGAGA 4754
D 1693 ATGACTTGGTTGAGTACTCACGAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAA 1752
Cp 4753 ATGACTTGGTTGAGTACTCACGAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAA 4694
D 1753 GAGAAATATGCAATGCTGTCATTAACCATAGTATTAATATGCTGCTGCTGCTGCTGCTG 1812
Cp 4693 GAGAAATATGCAATGCTGTCATTAACCATAGTATTAATATGCTGCTGCTGCTGCTGCTG 4634
D 1813 CAACGATCGGAGGACCGGAAGGAGCTTAACCGCTTTTTCACAAACATGGGGATCATGTAA 1872
Cp 4633 CAACGATCGGAGGACCGGAAGGAGCTTAACCGCTTTTTCACAAACATGGGGATCATGTAA 4574
D 1873 CTGCGCTTGGATGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGAGGCTGACA 1932
Cp 4573 CTGCGCTTGGATGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGAGGCTGACA 4514
D 1933 CCAGATGCTGTAGCAATGCGCAACGCTTTCGCAACATTAATTAATGCTGCTGCTGCTGCTG 1992
Cp 4513 CCAGATGCTGTAGCAATGCGCAACGCTTTCGCAACATTAATTAATGCTGCTGCTGCTGCTG 4454
D 1993 CTCTAGCTTCCCGGGCAACAAATTAATAGACTGATGAGGCGGATAAAGTTGAGGAGCCAC 2052
Cp 4453 CTCTAGCTTCCCGGGCAACAAATTAATAGACTGATGAGGCGGATAAAGTTGAGGAGCCAC 4394
D 2053 TTCTGGCTGCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATATAATCTGAGCGGCTGAGC 2112
Cp 4393 TTCTGGCTGCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATATAATCTGAGCGGCTGAGC 4334
D 2113 GTGGGTCTCGCGGTATCATTTGAGGACACTGCGGCGAGATGCTAAGGCTTCCGCTATCGTAG 2172
Cp 4333 GTGGGTCTCGCGGTATCATTTGAGGACACTGCGGCGAGATGCTAAGGCTTCCGCTATCGTAG 4274
D 2173 TTATCTACACGACCGGAGCTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGA 2232
Cp 4273 TTATCTACACGACCGGAGCTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGA 4214
D 2233 TAGTGCTCTACTGATTAAGCATTTGGTACTGTCACACCAAGTTTACTCATATATCTTT 2292
Cp 4213 TAGTGCTCTACTGATTAAGCATTTGGTACTGTCACACCAAGTTTACTCATATATCTTT 4154
D 2293 AGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTGTGATA 2352
Cp 4153 AGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTGTGATA 4084
D 2353 ATCTCATGACCAAAATCCCTTAACTGAGTTTTCGTTCCACTGAGCGTCAGACCCGCTAG 2412
Cp 4084 ATCTCATGACCAAAATCCCTTAACTGAGTTTTCGTTCCACTGAGCGTCAGACCCGCTAG 4034
D 2413 AAAAGATCAAGAGATCTTCTGAGATCTTTTCTGCGGCTAATCTGCTGCTGCTGCTGCTG 2472
Cp 4033 AAAAGATCAAGAGATCTTCTGAGATCTTTTCTGCGGCTAATCTGCTGCTGCTGCTGCTG 3974
D 2473 CAAAAAACCCACCCCTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2532

```


Cp 4993 GCTTCTCTGTTTTCCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGT 4934
Db 1513 TGGGTACAGAGTGGGTTTACATCGAATGATCTCAACAGCGGTAAAGATCCTTCAGAGTT 1572
Cp 4933 TGGGTACAGAGTGGGTTTACATCGAATGATCTCAACAGCGGTAAAGATCCTTCAGAGTT 4874
Db 1573 TTGGCCCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGG 1632
Cp 4873 TTGGCCCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGG 4814
Db 1633 TATTATCCGTTATGACGCGGCGGCAAGAGCAACTCGTGGCGGATACACTATTCTCAGA 1692
Cp 4813 TATTATCCGTTATGACGCGGCGGCAAGAGCAACTCGTGGCGGATACACTATTCTCAGA 4754
Db 1693 ATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCACTCTTACGGATGGCATGACAGTAA 1752
Cp 4753 ATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCACTCTTACGGATGGCATGACAGTAA 4694
Db 1753 GAGAATATGCACTGCGCATTAACCATGAGTGATAACACTTGGGCCAACTTACTCTGA 1812
Cp 4693 GAGAATATGCACTGCGCATTAACCATGAGTGATAACACTTGGGCCAACTTACTCTGA 4634
Db 1813 CAACGATCGGAGGACGGAAGGAGCTAACCGCTTTTTCACAACATGSGGATCATGTAA 1872
Cp 4633 CAACGATCGGAGGACGGAAGGAGCTAACCGCTTTTTCACAACATGSGGATCATGTAA 4574
Db 1873 CTCGCCCTTGATGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGGTGACA 1932
Cp 4573 CTCGCCCTTGATGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGGTGACA 4514
Db 1933 CCAGATGCGTGTAGCAATGGCAACAGCTTCGCGCAACTATTACTGGCGAATCTTTA 1992
Cp 4513 CCAGATGCGTGTAGCAATGGCAACAGCTTCGCGCAACTATTACTGGCGAATCTTTA 4454
Db 1993 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGGATGAAGTTTCAGAGCCAC 2052
Cp 4453 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGGATGAAGTTTCAGAGCCAC 4394
Db 2053 TTCTGCGCTCGGCCCTTCGCGCTGCGTGTATTAATGCTGATAAATCTGGAGCCGCTGAGC 2112
Cp 4393 TTCTGCGCTCGGCCCTTCGCGCTGCGTGTATTAATGCTGATAAATCTGGAGCCGCTGAGC 4334
Db 2113 GTGGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGGTAAGCCCTCCGCTATCGTAG 2172
Cp 4333 GTGGGTCTCGCGGTATCATTTGAGCACTGGGCGCAGATGGTAAGCCCTCCGCTATCGTAG 4274
Db 2173 TTATCTACAGCGCGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGA 2232
Cp 4273 TTATCTACAGCGCGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGA 4214
Db 2233 TAGTGCCCTCACTGATTAAAGCAATGGTAAGTGTGAGCAAGTTTACTCATATATCTTT 2292
Cp 4213 TAGTGCCCTCACTGATTAAAGCAATGGTAAGTGTGAGCAAGTTTACTCATATATCTTT 4154
Db 2293 AGATTGATTAAAGCAATCTTTTAAATTTAAAGGATCTAGTGAGATCCTTTTGTGATA 2352
Cp 4153 AGATTGATTAAAGCAATCTTTTAAATTTAAAGGATCTAGTGAGATCCTTTTGTGATA 4094
Db 2353 ATCTCATGACCAAAATCCCTTTAACGTGAGTTTTCGTTCCACTGAGCGGTACAGCCCGTAG 2412
Cp 4093 ATCTCATGACCAAAATCCCTTTAACGTGAGTTTTCGTTCCACTGAGCGGTACAGCCCGTAG 4034
Db 2413 AAAAGATCAAGAGATCTTCTTGAGATCCTTTTTCGCGGTAATCTGCTGCTTGCAAA 2472
Cp 4033 AAAAGATCAAGAGATCTTCTTGAGATCCTTTTTCGCGGTAATCTGCTGCTTGCAAA 3974
Db 2473 CAATAAACCCAGCTACAGCGGTGGTTTGTTCGCGGATCAAGAGTACCACTCTTT 2532
Cp 3973 CAATAAACCCAGCTACAGCGGTGGTTTGTTCGCGGATCAAGAGTACCACTCTTT 3914
Db 2533 TTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAACTCTGCTTCTAGTGAGC 2592
Cp 3913 TTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAACTCTGCTTCTAGTGAGC 3894

Db 2593 CGTAGTTAGGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTATA 2652
Cp 3853 CGTAGTTAGGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTATA 3794
Db 2653 TCCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTGTGTCTTACCGGTTTGACTCAA 2712
Cp 3793 TCCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTGTGTCTTACCGGTTTGACTCAA 3734
Db 2713 GACGATAGTTACCGATTAAGCGGAGCGCTGSGGCTCAAAGGCGGCTCTCTGTACNACG 2772
Cp 3733 GACGATAGTTACCGATTAAGCGGAGCGCTGSGGCTCAAAGGCGGCTCTCTGTACNACG 3674
Db 2773 CCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA 2832
Cp 3673 CCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA 3614
Db 2833 GCGCCACGCTTCCCGAAAGAGAGAGCGACAGTATCCGCTAAGGCGGAGGCTGAGGCTGAG 2892
Cp 3613 GCGCCACGCTTCCCGAAAGAGAGAGCGACAGTATCCGCTAAGGCGGAGGCTGAGGCTGAG 3554
Db 2893 CAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCGCTGGTATCTTTATAGCTCTGCG 2952
Cp 3553 CAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCGCTGGTATCTTTATAGCTCTGCG 3494
Db 2953 GGTTCGCCACCTCTGACTTGAGCGTTCGATTTTGTATGCTGCTGCTGCTGCTGCTGCT 3012
Cp 3493 GGTTCGCCACCTCTGACTTGAGCGTTCGATTTTGTATGCTGCTGCTGCTGCTGCTGCT 3434
Db 3013 TATGGAAGAACCCAGCAACGCGGCTTTTACGGTTCCTGGGCTTTTGTGCTGCTTTTG 3072
Cp 3433 TATGGAAGAACCCAGCAACGCGGCTTTTACGGTTCCTGGGCTTTTGTGCTGCTTTTG 3374
Db 3073 CTCATGTTCTTCTCGGTTATCCCTGATTCCTGATTAACCGCTATTACCGCTTTG 3132
Cp 3373 CTCATGTTCTTCTCGGTTATCCCTGATTCCTGATTAACCGCTATTACCGCTTTG 3314
Db 3133 AGTAGCTGATACCCGCTCGCGCAGCCGACGAGCGCAGCGAGTCACTAGTGAAGG 3192
Cp 3313 AGTAGCTGATACCCGCTCGCGCAGCCGACGAGCGCAGCGAGTCACTAGTGAAGG 3254
Db 3193 AAGCGAAGAGCGGCCAATAGCAAAACGCTCTCTCGCGGCTTTGCGGCTTCACTTAAT 3252
Cp 3253 AAGCGAAGAGCGGCCAATAGCAAAACGCTCTCTCGCGGCTTTGCGGCTTCACTTAAT 3194
Db 3253 GCAGCTGGCAGCACAGTTCCTCCGACTGAAAGCGGCGAGTGAGCGCAACTTAATG 3312
Cp 3193 GCAGCTGGCAGCACAGTTCCTCCGACTGAAAGCGGCGAGTGAGCGCAACTTAATG 3334
Db 3313 TGAGTTAGCTCACTATTAGGACCCCGAGGCTTTTACACTTTATGCTTCCGCGCTGATGT 3372
Cp 3133 TGAGTTAGCTCACTATTAGGACCCCGAGGCTTTTACACTTTATGCTTCCGCGCTGATGT 3074
Db 3373 TGTGTGAATTTGAGCGGATTAACAATTTTACACAGGAAACAGATGACCATGATACG 3432
Cp 3073 TGTGTGAATTTGAGCGGATTAACAATTTTACACAGGAAACAGATGACCATGATACG 3014
Db 3433 CCAAGCGGCAATTAACCTCTACTAAAGGAAACAAAGCTGGGTACCGGCTGCTGCTG 3492
Cp 3013 CCAAGCGGCAATTAACCTCTACTAAAGGAAACAAAGCTGGGTACCGGCTGCTGCTG 2954
Db 3493 AGGTGAC 3500
Cp 2953 AGGTGAC 2946

RESULT 8
LOCUS EV67875 10765 bp DNA circular SYN 16-DEC-1997
DEFINITION pESP-I yeast expression vector, complete sequence.
ACCESSION U67875
NID 92689262
KEYWORDS Expression vector pESP-I.
SOURCE

ORGANISM	Expression vector pESP-1 artificial sequence: expression vectors. 1 (bases 1 to 10765) Lu.Q., Bauer,J.C. and Greener,A. Using Schizosaccharomyces pombe as a host for expression and purification of eukaryotic proteins Gene 200 (1-2), 135-144 (1997) 98038484 2 (bases 1 to 10765) Lu.Q. Direct Submission Submitted (22-AUG-1996) pSP, Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA 3 (bases 1 to 10765) Grafsky,A. Direct Submission Submitted (16-DEC-1997) R&D, Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA Sequence update by submitter On Dec 16, 1997 this sequence version replaced gi:1527192. PESP-1 yeast expression vector is derived from plasmid from Stratagene for expression and purification of heterologous genes in Schizosaccharomyces pombe. Location/Qualifiers 1..10765 /organism="Expression vector pESP-1" /db_xref="taxon:51339" 1401..2261 /product="beta-lactamase" /db_xref="PID:q1527193" /translation="MSIQHFRVALPFFAAFCFLPVFAHPETLVKVKDAEDQLGARVGY IELDLNSGKILSEFRRPEPRMWSPTFKVLGGVAGFLGQVQLQISGLKQVLRVYSONDLVE YSPVTEKHITIDGMTPVELCSAITSQNTANILTTIGSPKELTAFLINMGDHTVRL DRNEPNEIAINPDRIETPVMATITLKGELLTASRQQLIDNMEADKVKAGLP LRSAIPAGFWIADKSGAGRSRGIIAALGPDGPKSRIVVIYITGSOATMDERNRQIA EIGASLIKHM" 5464..6558 /codon_start=1 /product="3-isopropylmalate dehydrogenase" /db_xref="PID:q269263" /translation="MSAPKKIVLPDGHVGGQETAEAIKVLKALSDVRNRVNFENH LIGGAIDAIQPLPDEALEASKKVDVALLGAGVGPWKGTGVSVRPEGLLKRKLQL YANLRPCNTAASLDLSPKIQAKGTDFVVRVELVGGYFGKRKEDGGVANDLVE QYVPEVQPIIRMAAFMALQHEPPLP IWSLDKANVLASRLRWKRTVEETIKNEFPTLK KQHLIDSAAMILGVKNPHLNGIITISNMGDIISDEASVPGSLGLLPSLASLAPLD VNTAFGLYEPCHSGAPDLPKVKNVNPATILISAAAMLKLSLNPESKAIEDAVKKVLID AGIPTGLGGSNSTTEVDAAVEKKILA" 8168..8884 /codon_start=1 /product="glutathione S-transferase" /db_xref="PID:q1527195" /translation="MSPILGWYKIGLVQPTKRLLEYLEEKYEHLVERDEGDKWRNK KFEGLFEFNLPVYIDGVKLTQSWAIIIRYIADKHNMLGGCPKPERAEISMLGKVALDI RYGVSRVAYSKQFETLKVDYLSKLPEMLKMFEDRCHTKYLNGDHTVHPDFMLYDALD VWLYNPMFLDAPFKLIVFPKKPIPAIPQIRKYLKSSKYIAMPVIGWQATFGGSGQHPPLK SDLVPGSGDTKDDKDGSPG"
BASE COUNT	3276 a 2125 c 2201 g 3163 t
ORIGIN	
Query Match	43.1%; Score 2228; DB 32; Length 10765.
Best local similarity 100.0%;	Prod No 0 00e-09;
Matches 2228; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db 1273	ATGGCACTTTTGGGGAAGATGCGGAGACCGCTATTGTTTATTTTCTAAATACATT 1332
Cp 5173	ATGGCACTTTTGGGGAAGATGCGGAGACCGCTATTGTTTATTTTCTAAATACATT 5114
Db 1333	CAAAATATGATCGCTATAGACAAATAACCCGATGAATAATGTTTCAATAATATCAAAA 1392
Cp 5113	CAAAATATGATCGCTATAGACAAATAACCCGATGAATAATGTTTCAATAATATCAAAA 5054
Db 1393	GAAGATATGATGATTTCAACATTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTT 1452

Cp 3973 CAAAAAACACCGCTACCGGCTGTTGTTGCGGATCAAGAGCTACCAACTCTTT 3914
 Db 2533 TTCGAGGTAACTGCTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTAGTGAGC 2592
 Cp 3913 TTCGAGGTAACTGCTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTAGTGAGC 3854
 Db 2593 CGTAGTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCTGCTGTA 2652
 Cp 3853 CGTAGTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCTGCTGTA 3794
 Db 2653 TCTGTATACAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2712
 Cp 3793 TCTGTATACAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3734
 Db 2713 GAGGATAGTTACCGGATAGGCGGCGAGCGGCTGCGGCTGAAAGCGGGGCTTCTGTCACACAGC 2772
 Cp 3733 GAGGATAGTTACCGGATAGGCGGCGAGCGGCTGCGGCTGAAAGCGGGGCTTCTGTCACACAGC 3674
 Db 2773 CAGGTTGGAGGAGGAGCTACACGGAAGTACAGGAGTACAGGAGTACAGGAGTACAGGAG 2832
 Cp 3673 CAGGTTGGAGGAGGAGCTACACGGAAGTACAGGAGTACAGGAGTACAGGAGTACAGGAG 3614
 Db 2833 GCGGAGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2892
 Cp 3613 GCGGAGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3554
 Db 2893 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2952
 Cp 3553 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3494
 Db 2953 GCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3012
 Cp 3493 GCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3434
 Db 3013 TATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3072
 Cp 3433 TATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3374
 Db 3073 CTCACATGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3132
 Cp 3373 CTCACATGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3314
 Db 3133 AGTGAGCTGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3192
 Cp 3113 AGTGAGCTGATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3254
 Db 3193 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3252
 Cp 3253 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3194
 Db 3253 GCAGCTGGCAGCAGAGGTTTCCGAGCTGGAAGCGGCGAGTACGAGCGCAACCAATTA 3312
 Cp 3193 GCAGCTGGCAGCAGAGGTTTCCGAGCTGGAAGCGGCGAGTACGAGCGCAACCAATTA 3334
 Db 3313 TGAGTTAGCTACCTATAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3372
 Cp 3133 TGAGTTAGCTACCTATAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3074
 Db 3373 TGTGTGAATTTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3432
 Cp 3073 TGTGTGAATTTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3014
 Db 3433 CCAAGCGGCAATTAAGCTTACCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3492
 Cp 3013 CCAAGCGGCAATTAAGCTTACCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2954
 Db 3493 AGGTCGAC 3500
 Cp 2953 AGGTCGAC 2946
 RESULT 9 SYNBLKSPV 2964 bp DNA circular SYN 04-JUN-1993
 LOCUS

DEFINITION
 BlueScribe KS Plus cloning vector.
 ACCESSION
 L08785
 NID
 g310729
 KEYWORDS
 Synthetic construct DNA.
 ORGANISM
 artificial construct
 1 (bases 1 to 2964)
 AUTHORS
 Gilbert, W.
 TITLE
 Obtained from VecBase 3.0
 JOURNAL
 Unpublished (1991)
 COMMENT
 These data and their annotation were supplied to GenBank by Will
 Gilbert under the auspices of the GenBank Curator Program.
 BlueScribe KS Plus - Cloning vector
 ENTRY BLUEKSP
 #TYPE DNA CIRCULAR
 TITLE BlueScribe KS Plus - Cloning vector
 DATE 28-JAN-1987

#sequence 02-FEB-1987
 #sequence 04-MAR-1987
 #sequence 03-APR-1987
 ACCESSION VB0078
 SOURCE artificial
 REFERENCE
 #number 1
 #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorac
 J.

#journal Gene (1987) in press
 #citation Sequence data from StrataGene
 #comment sequence correction according to StrataGene COMMENT
 Obtained from StrataGene on floppy disc.
 Revised 02-FEB-1987 by F. Pfeiffer:
 1409/10 'Ar' to 'TA' to match revised sequence of PBR322
 Revised 4-MAR-1987 to match sequence of pUC19 on request
 of StrataGene

Polylinker region revised 03-APR-1987 according to StrataGene
 COMMENT

The stand shown corresponds to pUC19c.
 As in the published sequence of pUC19c, The M13mp19 lacZ region
 is on the complementary strand.

COMMENT
 This vector contains the fl origin so that the plus strand
 can be obtained upon fl superinfection.

KEYWORDS
 CROSSREFERENCE

#parent

VecBase(3):BlueM13p

#parent

VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,

VecSource(3):bgalks, GenBank(50):PFI

#brother

VecBase(3):BlueKsm, VecBase(3):BlueSkp

PARENT

Features of BlueKsp (2964 bp)

residue source
 3- 458 5943-5488 (c) phage fl
 460- 524 236- 400 pUC19c
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKs-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-2964 448-2617 pUC19c
 Conflict (cfl) and Mutations (mut): none

PARENT

Features of BlueKsp (2964 bp)

residue source
 3- 458 5943-5488 (c) phage fl
 460- 624 449- 285 (c) pUC19
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKs-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-1031 237- 1 (c) pUC19
 1032-2964 2686- 754 (c) pUC19
 Conflict (cfl) and Mutations (mut): none
 FEATURE

643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-DraII-ApaI-XhoI-SalI-Clal-HindIII-EcoRV-EcoRI-PstI-
SmaI-BamHI-SpeI-XbaI-NciI-XmaII-BstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSP #length 2964 #checksum 690.
Location/Qualifiers
1. 2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 750 a 736 c 755 g 723 t
ORIGIN

Query Match 43.0% Score 2224; DB 32; Length 2964;
Best local Similarity 94.9% Pred. No. 6 60+00.
Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 734 CTCGACCTCGAGGGGGGGGGTACCCAGCTTTTGTTCCTTTTAGTGAGGGTTAATTCC 793
QY 2946 GTGGATCTGAGGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3005
Db 794 GAGCTTGGAGTAATCATGGTATAGTATGTTTGTCTGAGGAAATTTTATCGGCTCACAAT 853
QY 3006 GGGTTGGGTAATCATGGTATAGTATGTTTGTCTGAGGAAATTTTATCGGCTCACAAT 3065
Db 854 TCCACACATACAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913
QY 3066 TCCACACATACAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3125
Db 914 CTAACCTACATTAATTTGGGTCACATGGGCGCTTTCAGTCGGGAAACCGTGGTG 973
QY 3126 CTAACCTACATTAATTTGGGTCACATGGGCGCTTTCAGTCGGGAAACCGTGGTG 3185
Db 974 CGAGCTGCATTAATGAATGGGTAAGAGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 1033
QY 3186 CGAGCTGCATTAATGAATGGGTAAGAGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 3245
Db 1034 TTTGGTTCCTGCTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
QY 3246 TTTGGTTCCTGCTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305
Db 1094 AGCTCAGTCAAGAGGGGTAATACGCTTATCCAGAGATACAGGGGATACAGCAGAGAA 1153
QY 3306 AGCTCAGTCAAGAGGGGTAATACGCTTATCCAGAGATACAGGGGATACAGCAGAGAA 3365
Db 1154 CATGTGAGTCAAGAGGGTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1213
QY 3366 CATGTGAGTCAAGAGGGTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3425
Db 1214 TTTGATAGGCTCGGCTCGGCTGAGAGATACAGAGGATACAGAGGATACAGAGGAG 1273
QY 3426 TTTGATAGGCTCGGCTCGGCTGAGAGATACAGAGGATACAGAGGATACAGAGGAG 3485
Db 1274 GCGAAACCGGACAGGACTATAAGATATAGAGGATATAGAGGATATAGAGGATAT 1333
QY 3486 GCGAAACCGGACAGGACTATAAGATATAGAGGATATAGAGGATATAGAGGATAT 3545
Db 1334 CTCCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1393
QY 3546 CTCCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3605
Db 1394 CGTGGGCTTTCTCATAGTCAAGCTTAGGCTATGCTGCTGCTGCTGCTGCTGCTGCT 1453
QY 3606 CGTGGGCTTTCTCATAGTCAAGCTTAGGCTATGCTGCTGCTGCTGCTGCTGCTGCT 3665
Db 1454 CAAGCTGGAGAGTATGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1513
QY 3666 CAAGCTGGAGAGTATGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3725
Db 1514 CTATGCTCTTGGTCAAGCCGGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1573

QY 3726 CTATGCTCTTGGTCAAGCCGGTGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4785
Db 1574 TAACAGGATTAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1634
QY 3786 TAACAGGATTAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4845
Db 1634 TAACAGGCTTACATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1694
QY 3846 TAACAGGCTTACATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4905
Db 1694 CTTCGAAAAAAGAGTGGTACGCTTCATAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1754
QY 3906 CTTCGAAAAAAGAGTGGTACGCTTCATAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4965
Db 1754 TTTTTCGTTTGAAGCAGCAGAGATTACGGCGAGAAAAAAGAGGATCTCAGAGAGCTTT 1814
QY 3966 TTTTTCGTTTGAAGCAGCAGAGATTACGGCGAGAAAAAAGAGGATCTCAGAGAGCTTT 4975
Db 1814 GATCTTTTACGCGGGTCTGACGCTCAGTGGAAAGAGGAGGAGGAGGAGGAGGAGGAG 1874
QY 4026 GATCTTTTACGCGGGTCTGACGCTCAGTGGAAAGAGGAGGAGGAGGAGGAGGAGGAG 4085
Db 1874 CATGAGATTATCAAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAAATGAAGTTTAA 1934
QY 4086 CATGAGATTATCAAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAAATGAAGTTTAA 4145
Db 1934 ATCAATCTAAGATATATAGTAAAGCTTCTGCTGACAGGTTAGCAATGCTTATCACTCA 1994
QY 4146 ATCAATCTAAGATATATAGTAAAGCTTCTGCTGACAGGTTAGCAATGCTTATCACTCA 4205
Db 1994 GGCACCTATCTCAGCGATCTGCTATTTCCTTCATCCATAGTTCGCTGATCTGCTGCTG 2054
QY 4206 GGCACCTATCTCAGCGATCTGCTATTTCCTTCATCCATAGTTCGCTGATCTGCTGCTG 4265
Db 2054 GTAGATAACTACGATACGGAGGCTTACCATCTGCGCCGAGTCTGCTGATGATATATGAT 2114
QY 4266 GTAGATAACTACGATACGGAGGCTTACCATCTGCGCCGAGTCTGCTGATGATATATGAT 4275
Db 2114 AGACCCAGGCTCAGCGCTCCAGATTTATCAGCAATTAACAGAGGAGGAGGAGGAGGAG 2174
QY 4326 AGACCCAGGCTCAGCGCTCCAGATTTATCAGCAATTAACAGAGGAGGAGGAGGAGGAG 4385
Db 2174 GCGCAGAGTGGTCTGCTCAAGCTTTATCGGCTCAGTCCAGCTTATTAATTTTGGGAGGA 2234
QY 4386 GCGCAGAGTGGTCTGCTCAAGCTTTATCGGCTCAGTCCAGCTTATTAATTTTGGGAGGA 4445
Db 2234 AGCTAGAGTAGTAGTTCGCGAGCTTAATAGTTCGCGAGAGGAGGAGGAGGAGGAGGAG 2294
QY 4446 AGCTAGAGTAGTAGTTCGCGAGCTTAATAGTTCGCGAGAGGAGGAGGAGGAGGAGGAG 4505
Db 2294 CATGCTGGTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2354
QY 4506 CATGCTGGTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4565
Db 2354 AAGCGAGTTAT 2414
QY 4566 AAGCGAGTTAT 4625
Db 2414 GATGCTGTGAGAGTAGTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2474
QY 4626 GATGCTGTGAGAGTAGTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4685
Db 2474 TAATTCCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2534
QY 4686 TAATTCCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2594
Db 2534 CAAGCTGCATCTCAGAGTAGTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2594
QY 4746 CAAGCTGCATCTCAGAGTAGTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4805
Db 2594 GGTAAATACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2654

```

QY 4806 GGATAATACCGCCACATAGCAAACTTTAAAGTGTCTCATATTGAAACGTTCTTC 4865
Db 2654 GGGCGGAAAACCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG 2713
QY 4866 GGGGCGAAAACCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG 4925
Db 2714 TGACGCAACTGATCTTCAGCATCTTTTACTTCTACCGAGGTTCTGGGTGAGCAAAAAC 2773
QY 4926 GACACCAACTGATCTTCAGCATCTTTTACTTCTACCGAGGTTCTGGGTGAGCAAAAAC 4985
Db 2774 AGGAAGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCAT 2833
QY 4986 AGGAAGCAAAATGCGCAAAAAGGAATAAGGGCGACACGGAATGTTGAATACTCAT 5045
Db 2834 ACTCTTCCTTTTCAATATTATTAAGCACTTTATCATAGGTTTATGCTCATGAGCGGATA 2893
QY 5046 ACTCTTCCTTTTCAATATTATTAAGCACTTTATCATAGGTTTATGCTCATGAGCGGATA 5105
Db 2894 CATATTGAATGATTATAGAAAATTAACAATAGGGGTTCCGGCGACATTTCCCGGAAA 2953
QY 5106 CATATTGAATGATTATAGAAAATTAACAATAGGGGTTCCGGCGACATTTCCCGGAAA 5165
Db 2954 AGTGCCAC 2961
QY 5166 AGTGCCAC 5173

RESULT 10
LOCUS SYNBLKSMV 2964 bp DNA circular SYN 04-JUN-1993
DEFINITION Bluescribe KS Minus cloning vector.
ACCESSION L08784
NC_0310728
KEYWORDS Synthetic construct DNA.
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert W.
TITLE Obtained from Vecbase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program.
ENTRY BLUESKM
TITLE Bluescribe KS Minus - Cloning vector
DATE 28-JAN-1987
#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION V80077
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge
J.
#Journal Gene (1987) in press
#citation Sequence data from StrataGene
#comment sequence correction according to StrataGene COMMENT
Obtained from StrataGene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer:
1409/10 'AT' to 'TA' to match revised sequence of PBR322
Revised 4-MAR-1987 to match sequence of pUC19 on request
of StrataGene
Polylinker region revised 03-APR-1987 according to StrataGene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mp19 lacZ region
is on the complementary strand.
COMMENT
This vector contains the fl origin so that the minus strand
can be obtained upon fl superinfection.
KEYWORDS
CROSSREFERENCE

```

```

#parent
VecBase(3):BlueM13m
#parent
VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
VecSource(3):bGalkS, GenBank(50):PFI
#brother
VecBase(3):BlueKSp, VecBase(3):BlueSKm
PARENT
Features of BlueKSm (2964 bp)
residue source
3- 458 5488-5943 phage fl
460- 524 236- 400 pUC19c
526- 645 1- 20 17 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-2964 448-2617 pUC19c
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueKSm (2964 bp)
residue source
3- 458 5488-5943 phage fl
460- 524 449- 285 (c) pUC19
626- 645 1- 20 17 promoter
653- 760 108- 1 (c) BlueKS-polylinker
772- 791 20- 1 (c) T3 promoter
795-1031 237- 1 (c) pUC19
1032-2964 2686- 754 (c) pUC19
Conflict (cfl) and Mutations (mut): none
FEATURE 643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-DraII-ApaI-XhoI-SalI-HindIII-EcoRV-EcoRI-PstI-
SmaI-BamHI-SpeI-XbaI-NotI-XmaII-EstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSm #length 2964 #checksum 1589.
Location/Qualifiers
1. -2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 708 a 756 c 735 g 765 t
ORIGIN
Query Match 43.0% Score 2224; DB 32; Length 2964;
Best Local Similarity 95.9%; Pred. No. 0.00e+00;
Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 734 GTCGACCTCGAGGGGGGCGGTACCCAGCTTTTGTTCCTTTAGTCAGGGTTAAATTC 3005
QY 2946 GTCGACCTCGAGGGGGGCGGTACCCAGCTTTTGTTCCTTTAGTCAGGGTTAAATTC 3005
Db 794 GAGCTTGGCGTAAATCATGTGTCATAGCTGTTTCCTGTGTAATTTTATCGGCTCACAT 853
QY 3006 GCGCTTGGCGTAAATCATGTGTCATAGCTGTTTCCTGTGTAATTTTATCGGCTCACAT 3065
Db 854 TCCACACACATACAGCGCGGAGGATAAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAG 913
QY 3066 TCCACACACATACAGCGCGGAGGATAAAGTGTAAAGCTTGGGGTGCTTAATGAGTGAG 3125
Db 914 CTAACTCAGATAATTTGCGTTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973
QY 3126 CTAACTCAGATAATTTGCGTTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3185
Db 974 CCAGCTGCGATTAAATGAATCGGCAACGCGCGGAGAGAGCGGTTTGGCTATTTGGCGGTC 1033
QY 3186 CCAGCTGCGATTAAATGAATCGGCAACGCGCGGAGAGAGCGGTTTGGCTATTTGGCGGTC 3245
Db 1034 TTCGGTTCTCTGCTCAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
QY 3246 TTCGGTTCTCTGCTCAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305

```

DB 1094 AGCTCACTAAAGGGCGGTAAATAGCGTTATCCATACAACTACAGGATACAGCAGCAAGAA 1153
QY 3306 AGCTCACTAAAGGGCGGTAAATAGCGTTATCCATACAACTACAGGATACAGCAGCAAGAA 3365
DB 1154 CATCTGACGTAAGGAGGACGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1213
QY 3366 CATCTGACGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3425
DB 1214 TTTCGALAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1273
QY 4426 TTTCGALAGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3485
DB 1274 GCGAAATGCGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3333
QY 3486 GCGAAATGCGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3545
DB 1334 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3393
QY 3546 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3605
DB 1394 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1453
QY 4606 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3665
DB 1454 CAAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1513
QY 3666 CAAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3725
DB 1514 CTATCGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1573
QY 3726 CTATCGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3785
DB 1574 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1633
QY 3786 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3845
DB 1634 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1693
QY 3846 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3905
DB 1694 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1753
QY 4006 CTCTCGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3965
DB 1754 TTTTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1813
QY 4066 TTTTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4025
DB 1814 CATCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1873
QY 4026 CATCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4085
DB 1874 CATGAGATTATCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1933
QY 4086 CATGAGATTATCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4145
DB 1934 ATCAATGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1993
QY 4146 ATCAATGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4205
DB 1994 GCGAG 2053
QY 4206 GCGAG 4265
DB 2054 GTATGAG 2113
QY 4266 GTATGAG 4325
DB 2114 AGAG 2173
QY 4326 AGAG 4385
DB 2174 GCGAG 2233

QY 4386 GCGAG 4444
DB 2234 AGCTCACTAAAGGGCGGTAAATAGCGTTATCCATACAACTACAGGATACAGCAGCAAGAA 2293
QY 4446 AGCTCACTAAAGGGCGGTAAATAGCGTTATCCATACAACTACAGGATACAGCAGCAAGAA 4505
DB 2294 CATCTGAG 2353
QY 4506 CATCTGAG 4565
DB 2354 AAGGAG 2413
QY 4566 AAGGAG 4625
DB 2414 GAG 2473
QY 4626 GAG 4685
DB 2474 TAACTGAG 2533
QY 4686 TAACTGAG 4745
DB 2534 CAAGTCATCTGAG 2593
QY 4746 CAAGTCATCTGAG 4651
DB 2594 GATATAAGGAG 2653
QY 4806 GATATAAGGAG 4865
DB 2654 GCGGAG 2713
QY 4866 GCGGAG 4925
DB 2714 TGCAGGAG 2773
QY 4926 TGCAGGAG 4991
DB 2774 AGGAG 2833
QY 4986 AGGAG 5045
DB 2834 ATCTGAG 2903
QY 5046 ATCTGAG 3061
DB 2894 CATATGAG 2953
QY 5106 CATATGAG 3011
DB 2954 AGTGGCAG 2961
QY 5166 AGTGGCAG 5173

RESULT 11
LOCUS 002449 2967 bp DNA circular SYN 29-MAR-1999
DEFINITION Cloning vector pDIRECT, complete sequence.
ACCESSION U02449
NID 9413815
KEYWORDS Cloning vector pDIRECT.
ORGANISM Cloning vector pDIRECT.
REFERENCE Cloning vector pDIRECT.
AUTHORS 1 (bases 1 to 2967)
TITLE Cloning vector pDIRECT.
JOURNAL Cloning vector pDIRECT.
AUTHORS 2 (bases 1 to 2967)
TITLE Cloning vector pDIRECT.
JOURNAL Cloning vector pDIRECT.
AUTHORS 2 (bases 1 to 2967)
TITLE Cloning vector pDIRECT.
JOURNAL Cloning vector pDIRECT.

1020 East Meadow Circle, Palo Alto, CA 94303, USA
 This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3.
 This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH; this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail: TECH@CLONTECH.COM.

FEATURES Location/Qualifiers
 source 1..2967
 /organism="Cloning vector pDIRECT"
 /db_xref="taxon:31799"
 BASE COUNT 726 a 756 c 735 g 750 t
 ORIGIN
 Query Match 43.0% Score 2224; DB 32; Length 2967;
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;
 Matches 2224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 700 GGGGACATTTTCGGGAAATGTGGGAAACCCCTATTGTTTATTTTCTAAATACATT 759
 Cp 5173 GTGGACATTTTCGGGAAATGTGGGAAACCCCTATTGTTTATTTTCTAAATACATT 5114
 Db 760 CAAATATGATCCGCTCATGAGCAATAAACCTGATAAATGCTTCATATATTGAAAA 819
 Cp 5113 CAAATATGATCCGCTCATGAGCAATAAACCTGATAAATGCTTCATATATTGAAAA 5054
 Db 820 GCAAGATGATGATGATCAACATTCGGTGTGGCCCTATTCCCTTTTTCGGGCATTTT 879
 Cp 5053 GCAAGATGATGATGATCAACATTCGGTGTGGCCCTATTCCCTTTTTCGGGCATTTT 4994
 Db 880 GCTTCCTGTTTTTTCCTACCCAGAAAGCGTGTGAAAGTAAAGATGCTGAAGATCAGT 939
 Cp 4993 GCTTCCTGTTTTTTCCTACCCAGAAAGCGTGTGAAAGTAAAGATGCTGAAGATCAGT 4934
 Db 940 TGGGTGCAGAGTGGGTTTACATCGAAGTATGATCTCAACAGCGGTAGATCCTTGAGATT 999
 Cp 4933 TGGGTGCAGAGTGGGTTTACATCGAAGTATGATCTCAACAGCGGTAGATCCTTGAGATT 4874
 Db 1000 TTCGCTCCGGAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG 1059
 Cp 4873 TTCGCTCCGGAAGAACTTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG 4814
 Db 1060 TATTATCCGCTATTACGCGCGGGAAGAGCAACTGCTGCTGCTGCTATACACTATTCTCAGA 1119
 Cp 4813 TATTATCCGCTATTACGCGCGGGAAGAGCAACTGCTGCTGCTGCTATACACTATTCTCAGA 4754
 Db 1120 ATGATTTGGTTGAGTACTACAGTTCAGAAAGCAATCTTACGGATGGCATCAGATGA 1179
 Cp 4753 ATGATTTGGTTGAGTACTACAGTTCAGAAAGCAATCTTACGGATGGCATCAGATGA 4694
 Db 1180 GAGAATTATGAGTCTGCCATACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
 Cp 4593 GAGAATTATGAGTCTGCCATACCATGATGATGATGATGATGATGATGATGATGATGATGAT 4634
 Db 1240 CAACGATCGGAGACCGAAGAGCTTAACCGGTTTTTTCACAACTATGCGGCAACTTACTTCTGA 1299
 Cp 4633 CAACGATCGGAGACCGAAGAGCTTAACCGGTTTTTTCACAACTATGCGGCAACTTACTTCTGA 4574
 Db 1300 CTCGGTTGATGTTGGGAAAGCGAGCTGAATGAAGCCATACCAAGAGCGAGCGTGACA 1359
 Cp 4573 CTCGGTTGATGTTGGGAAAGCGAGCTGAATGAAGCCATACCAAGAGCGAGCGTGACA 4514
 Db 1360 CCACATGCTCTAGCAATGSCAACCAAGCTTGCACAACTATTAACTGGGCAACTACTTA 1419
 Cp 4513 CCACATGCTCTAGCAATGSCAACCAAGCTTGCACAACTATTAACTGGGCAACTACTTA 4454

Db 1420 CTCTAGCTTCCCGGACACATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
 Cp 4453 CTCTAGCTTCCCGGACACATTAATAGATGATGATGATGATGATGATGATGATGATGATGAT 4394
 Db 1480 TTCTGCTCTCGGCTCTTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
 Cp 4393 TTCTGCTCTCGGCTCTTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4334
 Db 1540 GTGGTCTCTCGGCTCTTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
 Cp 4333 GTGGTCTCTCGGCTCTTTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4274
 Db 1600 TTATCTACAGACGCGGAGTTCAGGCAACTATGATGATGATGATGATGATGATGATGATGATGAT 1659
 Cp 4273 TTATCTACAGACGCGGAGTTCAGGCAACTATGATGATGATGATGATGATGATGATGATGATGAT 4214
 Db 1660 TAGTGCTCTCACTGATTAAGCAATTTTAAATTTAAAGGATCTAGGTAAGATCTTCTTCTGATA 1719
 Cp 4213 TAGTGCTCTCACTGATTAAGCAATTTTAAATTTAAAGGATCTAGGTAAGATCTTCTTCTGATA 4154
 Db 1720 AGATTGATTTAAACTTTCATTTTAAATTTAAAGGATCTAGGTAAGATCTTCTTCTGATA 1779
 Cp 4153 AGATTGATTTAAACTTTCATTTTAAATTTAAAGGATCTAGGTAAGATCTTCTTCTGATA 4094
 Db 1780 ATCTCATGACCAAAATCCCTTAACTGAGTTCCTTCCACTGAGCGCTGAGACCCCTAG 1839
 Cp 4093 ATCTCATGACCAAAATCCCTTAACTGAGTTCCTTCCACTGAGCGTCAAGCGCCCTAG 4034
 Db 1840 AAAAGATCAAGGATCTTCTTGAATCTTCTTGAATCTTCTTGAATCTTCTTGAATCTTCTTGA 1899
 Cp 4033 AAAAGATCAAGGATCTTCTTGAATCTTCTTGAATCTTCTTGAATCTTCTTGAATCTTCTTGA 3974
 Db 1900 CAAAAAACCACCGCTACAGCGGCTTTCCTGCGGATCAAGAGCTACCAACTCTTT 1959
 Cp 3973 CAAAAAACCACCGCTACAGCGGCTTTCCTGCGGATCAAGAGCTACCAACTCTTT 3914
 Db 1960 TTCGAAAGTAACTGGCTTCAGCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2019
 Cp 3913 TTCGAAAGTAACTGGCTTCAGCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3854
 Db 2020 GGTAGTTAGCCACCATCTTCAAGAACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2079
 Cp 3853 GGTAGTTAGCCACCATCTTCAAGAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3794
 Db 2080 TCTGTTTACAGTGGCTGCTGCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2139
 Cp 3793 TCTGTTTACAGTGGCTGCTGCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3734
 Db 2140 GACGATAGTTACCGGATGAGCGGCTGCGGCTGAGCGGGGGIICGICGACACAGC 2199
 Cp 3733 GACGATAGTTACCGGATGAGCGGCTGCGGCTGAGCGGGGGIICGICGACACAGC 3674
 Db 2200 CCAGCTTGGAGCAACCATCTACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2259
 Cp 3673 CCAGCTTGGAGCAACCATCTACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3614
 Db 2260 CGCGCAGCTTCCCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2319
 Cp 3613 CGCGCAGCTTCCCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3554
 Db 2320 CAGGAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2374
 Cp 3553 CAGGAGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3494
 Db 2380 GGTTCGCGCTCTGACCTGAGCGTCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2439
 Cp 3493 GGTTCGCGCTCTGACCTGAGCGTCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3434
 Db 2440 TATGAAAAAGCCAGCAACCGGCTTTTATGAGGTTCTTGGGCTTTTGTGCGCTTTTG 2499
 Cp 3433 TATGAAAAAGCCAGCAACCGGCTTTTATGAGGTTCTTGGGCTTTTGTGCGCTTTTG 3374
 Db 2500 CTCATGCTCTTCTTCTGCGCTTATCCCTCTGATCTGTGGGATAACCGCTTATACCGCTTTG 2559

Db 1685 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTTCGAGGACCC 1744
 Cp 4453 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTTCGAGGACCC 4394
 Db 1745 TTCTGGGCTCGGCGCTTCCGGGTGGCTGGTTTATTGCTGATAAATCTCGAGCCGGTGAAG 1804
 Cp 4393 TTCTGGGCTCGGCGCTTCCGGGTGGCTGGTTTATTGCTGATAAATCTCGAGCCGGTGAAG 4334
 Db 1805 GTGGGCTCGGCTGATCAATTCAGAGCACTGGGCGCAGATGGTAAGCCCTCCCGTATCGTAG 1864
 Cp 4333 GTGGGCTCGGCTGATCAATTCAGAGCACTGGGCGCAGATGGTAAGCCCTCCCGTATCGTAG 4274
 Db 1865 TTATCTACAGAGCGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA 1924
 Cp 4273 TTATCTACAGAGCGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGA 4214
 Db 1925 TAGTGCTCAGCTGATTAAGCATTTGGTAAGTGTGACAGCAAGTTTACTCATATATATCTTT 1984
 Cp 4213 TAGTGCTCAGCTGATTAAGCATTTGGTAAGTGTGACAGCAAGTTTACTCATATATATCTTT 4154
 Db 1985 AGATTGATTTAAAGCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTTGATA 2044
 Cp 4153 AGATTGATTTAAAGCTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCTCTTTTGATA 4094
 Db 2045 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAG 2104
 Cp 4093 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAG 4034
 Db 2105 AAAAGATCAAGATCTCTTTGAGATCTCTTTTCTGGCGGTATCTGCTGCTTGCAAA 2164
 Cp 4033 AAAAGATCAAGATCTCTTTGAGATCTCTTTTCTGGCGGTATCTGCTGCTTGCAAA 3974
 Db 2165 CAAAAAACCACCGCTACAGCGGTGTTTGTGGCGGATCAAGAGTACCAACTCTTT 2224
 Cp 3973 CAAAAAACCACCGCTACAGCGGTGTTTGTGGCGGATCAAGAGTACCAACTCTTT 3914
 Db 2225 TTCCGAAGTAACTGGCTTTCAGCAGAGCGAGATACCAAACTACTGCTTCTAGTGTAGC 2284
 Cp 3913 TTCCGAAGTAACTGGCTTTCAGCAGAGCGAGATACCAAACTACTGCTTCTAGTGTAGC 3854
 Db 2285 GGTAGTTAGCGGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTGCTAA 2344
 Cp 3853 GGTAGTTAGCGGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCTGCTAA 3794
 Db 2345 TCTGTACTAGCTGGTGTGCTGAGTGGGATAAAGTGTGTTTACCGGTTGGACTCAA 2404
 Cp 3793 TCTGTACTAGCTGGTGTGCTGAGTGGGATAAAGTGTGTTTACCGGTTGGACTCAA 3734
 Db 2405 GAGATAGTTACCGGATAGCGGCTACCGCACTGAGATACCTACAGCGTGGGTGCTGACACAGC 2464
 Cp 3733 GAGATAGTTACCGGATAGCGGCTACCGCACTGAGATACCTACAGCGTGGGTGCTGACACAGC 3674
 Db 2465 CCAGCTTGGAGCAAGCGACCTACCGCACTGAGATACCTACAGCGTGGGTGCTGACACAGC 2524
 Cp 3673 CCAGCTTGGAGCAAGCGACCTACCGCACTGAGATACCTACAGCGTGGGTGCTGACACAGC 3614
 Db 2525 GCGGACGCTTCCCGAAGGAGAAAGCGCAGATATCCGGTAAAGCGGAGGTCGAA 2584
 Cp 3613 GCGGACGCTTCCCGAAGGAGAAAGCGCAGATATCCGGTAAAGCGGAGGTCGAA 3554
 Db 2585 CAGGAGCGCAGAGGAGCTTCCAGGGGGAACCGCTGATCTTTATAGTCTGCTG 2644
 Cp 3553 CAGGAGCGCAGAGGAGCTTCCAGGGGGAACCGCTGATCTTTATAGTCTGCTG 3494
 Db 2645 GGTTCGCACTCTGACTTACGCTGCTGATTTTGTGATGCTGCTGAGGCGGAGCC 2704
 Cp 3493 GGTTCGCACTCTGACTTACGCTGCTGATTTTGTGATGCTGCTGAGGCGGAGCC 3434
 Db 2705 TATGAAAAACGCGACGAGCGGCTTTTATAGGTTCTGCGCTTTTGTGCGCTTTG 2764
 Cp 3433 TATGAAAAACGCGACGAGCGGCTTTTATAGGTTCTGCGCTTTTGTGCGCTTTG 3374

RESULT 13
 LOCUS A58522 7474 bp DNA circular PAT 05-MAR-1998
 DEFINITION Sequence 26 from Patent WO9638555.
 ACCESSION A58522
 NID g3714138
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 7474)
 AUTHORS Bogart, T., Stringham, E. and Vandekerckhove, J.
 TITLE PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR USE IN THE CONTROL OF CELL BEHAVIOUR
 JOURNAL Patent: WO 9638555-A 26 05-DEC-1996;
 COMMENT BOGART THIERRY (BE)
 FEATURES Other publication AU 6123496 961218.
 source Location/Qualifiers
 1. 7474
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 2107 a 1921 c 1656 g 1790 t
 ORIGIN
 Query Match 43.0% Score 2222; DB 22; Length 7474.
 Best Local Similarity 100.0%; Pred No. 0.00e+00;
 Matches 2222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 5253 CTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTCGCGCTT 5312
 Qy 2952 CTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTGAGGGTTAAATTCGCGCTT 3011
 Db 5313 GCGGTAATCATGTCATAGCTTTTCCCTGTGTAATTTTATCGGCTCAATTCACCA 5372
 Qy 3012 GCGGTAATCATGTCATAGCTTTTCCCTGTGTAATTTTATCGGCTCAATTCACCA 3071
 Db 5373 CAACATACAGAGCGGAGGATAAAGTGTAAAGCTTGGGTCCTTAAGTGTAGTAACT 5432
 Qy 3072 CAACATACAGAGCGGAGGATAAAGTGTAAAGCTTGGGTCCTTAAGTGTAGTAACT 3131

KEYWORDS

SOURCE Cloning vector pFUN.
 ORGANISM Cloning vector pFUN
 artificial sequence; cloning vectors.
 REFERENCE 1 (bases 1 to 8072)
 AUTHORS Poquet, I., Ehrlich, S.D. and Gruss, A.
 TITLE An export-specific reporter designed for gram-positive bacteria:
 application to *Lactococcus lactis*
 J. Bacteriol. 180 (7), 1904-1912 (1998)
 JOURNAL 98196737
 MEDLINE 2 (bases 1 to 8072)
 REFERENCE Poquet, I. and Gruss, A.
 AUTHORS Direct Submission
 TITLE Submitted (15-DEC-1997) Laboratoire de Genetique Appliquee-URLGA,
 JOURNAL Institut National de la Recherche Agronomique, CRJ, Jouy en Josas
 78352, France

FEATURES

Location/Qualifiers
 1..8072
 /organism="Cloning vector pFUN"
 /db_xref="taxon:74538"
 /note="plasmid vector that creates translational fusions
 with a nuclease reporter without a signal peptide"
 misc_feature 1..4580
 /note="pL252 moiety; similar to sequence deposited under
 GenBank Accession Number AF039139"
 misc_feature 4581..7472
 /note="similar to BlueScript"
 terminator 7473..7507
 /note="trpA terminator"
 misc_feature 7512..7547
 /note="multiple cloning site"
 gene <7543..8010
 /note="truncated nuc gene potentially encoding deltaSP-Nuc
 ORF"
 CDS
 /gene="delta-nuc"
 <7543..8010
 /gene="delta-nuc"
 /note="nuclease from *Staphylococcus aureus* devoid of its
 signal peptide and the thirteen amino acids at the
 N-terminus of mature protein"
 /codon_start=1
 /transl_table=11
 /product="deltaSP-Nuc"
 /db_xref="PID:33043925"
 /translation="DPTVYSATSKKLHKEPATLIKIDGTVKLMYKGPMTFRLL
 VDPETKHKGVKVEYGEASFTKKMVENAKIEVEFDKQRTDKYGRGLAVYADG
 KWNVALVROGLKAVYVKPNNTHEQHLRSEAKKEKLNINSEADNSGQ"
 BASE COUNT 2717 a 1474 c 1694 g 2187 t

Query Match 43.0% Score 2234; DB 32; Length 8072;
 Best local similarity 99.9% Pred. No. 0.00e+00;
 Matches 2226; Conservative 0. Mismatches 2. Indels 0; Gaps 0;


Db 5240 GTGGCACTTTTCGGGNAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATT 5299
 Cp 5173 GTGGCACTTTTCGGGNAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATT 5114
 Db 5300 CAATATGATATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATTCAAAA 5359
 Cp 5113 CAATATGATATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATTCAAAA 5054
 Db 5360 GGAAGATGATGATTAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGGCATTTT 5419
 Cp 5053 GGAAGATGATGATTAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGGCATTTT 4994
 Db 5420 GCTTTCCTGTTTTGTTCACCCAGAACCGTGTGGAAGTAAAGATGCTGAAGATCAGT 5479
 Cp 4993 GCTTTCCTGTTTTGTTCACCCAGAACCGTGTGGAAGTAAAGATGCTGAAGATCAGT 4934
 Db 5480 TGGGTGCACGAGTGGGTGTACATCGAACTGGATCTCAACAGCGGTGAAGATCCTTTGAGGTT 5539
 Cp 4933 TGGGTGCACGAGTGGGTGTACATCGAACTGGATCTCAACAGCGGTGAAGATCCTTTGAGGTT 4874

Db 5540 TTCGCGCGGAAGACGTTTTTCAATGATGATGACACTTTTAAAGTTTCTGCTATGTCGCGCG 5599
 Cp 4873 TTCGCGCGGAAGACGTTTTTCAATGATGATGACACTTTTAAAGTTTCTGCTATGTCGCGCG 4814
 Db 5600 TATTATCCCGTATTACGCCCGGCAAGACAACTCGGTGCGGCATACATAATTCTCAGA 5659
 Cp 4813 TATTATCCCGTATTACGCCCGGCAAGACAACTCGGTGCGGCATACATAATTCTCAGA 4754
 Db 5660 ATGACTTGGTTGACTACTCACCAGTCACAGAAAGCATCTTTACGGATGGCATGACATAA 5719
 Cp 4753 ATGACTTGGTTGACTACTCACCAGTCACAGAAAGCATCTTTACGGATGGCATGACATAA 4694
 Db 5720 GAGAATATGCACTGCTGCGCATTAACCATGATGAACACTGCGGCCAACTTACTTTTGA 5779
 Cp 4693 GAGAATATGCACTGCTGCGCATTAACCATGATGAACACTGCGGCCAACTTACTTTTGA 4634
 Db 5780 CAACGATCGGAGCGGCAAGGAGCTAACCGTTTTTTGCACAACTATTAACCTGGCCAACTTAA 5839
 Cp 4633 CAACGATCGGAGCGGCAAGGAGCTAACCGTTTTTTGCACAACTATTAACCTGGCCAACTTAA 4574
 Db 5840 CTCGCTTCATCGTTGGGAACCGGAGCTGAATGAAGCATACCAAGCATACCAAGCAGAGCGI 5899
 Cp 4573 CTCGCTTCATCGTTGGGAACCGGAGCTGAATGAAGCATACCAAGCATACCAAGCAGAGCGI 4514
 Db 5900 CCACGATGCTGTAGCAATGCGCAACAGCTTGGCCAACTATTAACCTGGCCAACTTAA 5959
 Cp 4513 CCACGATGCTGTAGCAATGCGCAACAGCTTGGCCAACTATTAACCTGGCCAACTTAA 4434
 Db 5960 CTCAGCTCCCGGCAACAAITTAATAGACTGATGAGAGGAGGAGATAAAGTTGACAGAGACAC 6019
 Cp 4453 CTCAGCTCCCGGCAACAAITTAATAGACTGATGAGAGGAGGAGATAAAGTTGACAGAGACAC 4394
 Db 6020 TTCTGCTCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6079
 Cp 4393 TTCTGCTCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4334
 Db 6080 GTGGGTCTCGCGCTTCATTTGAGCACTGCGGCGCAGATGGTAAAGCTTCGCGCTATCTGAG 6139
 Cp 4333 GTGGGTCTCGCGCTTCATTTGAGCACTGCGGCGCAGATGGTAAAGCTTCGCGCTATCTGAG 4274
 Db 6140 TTATCTACACGACGGGAGTCAAGCACTATGGATGAACAGAAATAGACAGATCGCTCAGA 6199
 Cp 4273 TTATCTACACGACGGGAGTCAAGCACTATGGATGAACAGAAATAGACAGATCGCTCAGA 4214
 Db 6200 TAGTGCCTCAGCTGATTAAGCACTTGGTAACCTCAGACCAAGTTTACTCATATATACATT 6259
 Cp 4213 TAGTGCCTCAGCTGATTAAGCACTTGGTAACCTCAGACCAAGTTTACTCATATATACATT 4154
 Db 6260 AGATTGATTTAAACCTTCATTTTAAATTAAGGATCTAGGTGAAGATCGCTTTTGTGATA 6319
 Cp 4153 AGATTGATTTAAACCTTCATTTTAAATTAAGGATCTAGGTGAAGATCGCTTTTGTGATA 4094
 Db 6320 ATCTCATGACCAAAATCCCTTAACGTGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6379
 Cp 4093 ATCTCATGACCAAAATCCCTTAACGTGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4034
 Db 6380 AAAAGATCAAGAGATCTCTTCAGATCCCTTTTCTCGGCTATCTGCTGCTGCTGCTGCTGCTGCT 6439
 Cp 4033 AAAAGATCAAGAGATCTCTTCAGATCCCTTTTCTCGGCTATCTGCTGCTGCTGCTGCTGCTGCT 3974
 Db 6440 CAAAAAACCCCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTT 6499
 Cp 3973 CAAAAAACCCCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTT 3914
 Db 6500 TTCGGAAGTAACTGGCTTCACAGAGCGCAGATACCAAACTACTGCTTCTAGTGTAGC 6559
 Cp 3913 TTCGGAAGTAACTGGCTTCACAGAGCGCAGATACCAAACTACTGCTTCTAGTGTAGC 3854
 Db 6560 CGTAGTTAGCGCACCACTCAAGAACTCTGTAGACCGCGCTACATACCTCGCTCTGCTTAA 6619
 Cp 3853 CGTAGTTAGCGCACCACTCAAGAACTCTGTAGACCGCGCTACATACCTCGCTCTGCTTAA 3784

Db	2534	CAAGTCATTCTGGAATAGTGTATGGCGGACCGAGTTGCTCTTTGGCCGGGGTCAATACG	2593
Qy	4746	CAAGTCATTCTGGAATAGTGTATGGCGGACCGAGTTGCTCTTTGGCCGGGGTCAATACG	4805
Db	2594	GGATAATACGGCGCCACATAGCAGAACTTTAAAGTGTCTCAICATTTGAAAAAGTTTTC	2653
Qy	4806	GGATAATACGGCGCCACATAGCAGAACTTTAAAGTGTCTCAICATTTGAAAAAGTTTTC	4865
Db	2654	GGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCG	2713
Qy	4866	GGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCG	4925
Db	2714	TGCACCCAACTGATCTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGTGAGCAAAAAC	2773
Qy	4926	TGCACCCAACTGATCTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGTGAGCAAAAAC	4985
Db	2774	AGGAAGCAAAATTCGCGCAAAAAGGGAATAAGCGCGACACGGAAATGTTGAATACTCAT	2833
Qy	4986	AGGAAGCAAAATTCGCGCAAAAAGGGAATAAGCGCGACACGGAAATGTTGAATACTCAT	5045
Db	2834	ACTCTCCCTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATA	2893
Qy	5046	ACTCTCCCTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATA	5105
Db	2894	CATATTGGAATGTATTAGAAAAATAACAAATAGGGGTTCCGGGACATTTCCCGGAAA	2953
Qy	5106	CATATTGGAATGTATTAGAAAAATAACAAATAGGGGTTCCGGGACATTTCCCGGAAA	5165
Db	2954	AGTGC 2958	
Qy	5166	AGTGC 5170	

Search completed: Tue Mar 16 09:48:54 1999
Job time : 13754 secs.





(TM)

Release 3 1A John F Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

wpsrch_nn n a - n a database search, using Smith-Waterman algorithm
Run on: Tue Mar 16 12:05:21 1999; MasPar time 731 94 Seconds
          961.559 Million cell updates/sec
          tabular output not generated.

```

```

Title:
Description:
Perfect Score:
N.A. Sequence:
Comp
>US-09-020-716-7
(1-5173) from US09020716.seq
5173
1 CTAATGTGTAACGGTTAATA...ATTCCCGGAAAGATGCCAC 5173
GATTTAACATTGGCAATTAI...TAAAGGGCTTTTCACAGTC

```

Scoring table. TABLE default
Gap 5

Nmatch STD : Dbase 0: Query 0

Searched: 188442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

```
Database:
n-genseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part22 22:part23 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part36 36:part36 37:part37 38:part38
39:part39 40:part40
```

Statistics: Mean 10.428, Variance 6.781, scale 1.538

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match	%					
1	2222	43.0	5178	27	749876		pret-Splice.	0.00e+00
2	2222	43.0	7474	31	771350		Flasmiid pIB73 encodin	0.30e+00
3	2220	42.9	6206	27	749877		Autoregulatory vector	0.00e+00
4	2202	42.6	4145	7	040279		Sequence of clone pS2	0.00e+00
5	2203	42.6	5314	3	013576		Plasmiid pMTV1 contain	0.00e+00
6	2202	42.6	5356	25	743794		Plasmiid pRPHAT (rat	0.00e+00
7	2202	42.6	5534	25	743137		pMIOGII sequence incl	0.00e+00
8	2201	42.5	3681	2	013578		Plasmiid pKDEL5.	0.00e+00
9	2198	42.5	3699	39	V14340		Plasmiid pBSGFP expres	0.00e+00
10	2198	42.5	4277	7	040280		Sequence of clone pS2	0.00e+00
11	2196	42.5	4539	14	087347		Plasmiid pMVI.	0.00e+00
12	2200	42.5	4792	15	084696		Plasmiid GS contig. N.	0.00e+00
13	2200	42.5	5042	15	084694		Plasmiid glucoamylase	0.09e+00

Sequence of clone p42	0.00e+00
Plasmid pWR33169 enc0	0.00e+00
Plasmid pm16-1.	0.00e+00
Plasmid pm16.	0.00e+00
Plasmid pSVFL(+)	0.00e+00
Plasmid pSVFL(+)	0.00e+00
Vector p37M-100.	0.00e+00
Construct pGEM-HTR(KA	0.00e+00
Construct pGEM-HTR CO	0.00e+00
Plasmid pmP6-112.	0.00e+00
Plasmid pWR33196 enc0	0.00e+00
Plasmid pCRM1.8 conta	0.00e+00
Human cyclin D1-human	0.00e+00
Shortened C1 gene in	0.00e+00
pBEL10.	0.00e+00
PrE110; plasmid DNA r	0.00e+00
pBEL08.	0.00e+00
Interleukin-2 express	0.00e+00
pVE108.	0.00e+00
Plasmid pVE108 used i	0.00e+00
pE4; plasmid DNA rep	0.00e+00
pRLD12D3.Apal constr	0.00e+00
Plasmid pVE144 used i	0.00e+00
IL-4.Y124D/IqgI fusio	0.00e+00
Sequence of pT64.	0.00e+00
Plasmid pHCV-167 codi	0.00e+00
pHCV167 sequence	0.00e+00
Plasmid pHCV-162 codi	0.00e+00
Plasmid pCB51 encodin	0.00e+00
Plasmid pT62 encodin	0.00e+00
Plasmid pCB50 encodin	0.00e+00
Alphavirus-based euka	0.00e+00

ALIGNMENTS

RESULT 1
ID T49876 standard: DNA: 5178 BP.

```
T49876;      28-MAR-1997   (first entry)
DE           pret-Splice.
DI           pret-Splice.
DD           gene expression; transgenic animal; animal model; drug screening.
KW           vector, ds.
WW           Synthetic.
CS           Key
KS           Location/Qualifiers
FT           complement(2591)
              /tag= a
FT           /note="putative start site of trxn"
FT           tata_signal    complement(2616..2622)
              /*tag= b
FT           W09640946-Al.
PD           19-DEC-1996.
PE           07-JUN-1986; U10109.
PF           07-JUN-1995; US-474169.
PI           (UYA ) UNIV YALE.
PT           Schatz DG;
PR           WPI: 97-077273/07.
PS           Nucleic acid encoding tetracycline transactivator fusion protein -
             provides rapid and reversible control of gene expression, e.g. for
             creating animal models for drug screening
             Example 1; Fig 9b-g; 82pp: English.
CC           Vector preT-splice (r49876) is utilised in the construction of
             autoregulatory vector plasmid pTet-trak (r49877). Splice-pA was
             made by ligating the SV40 small T antigen intervening sequence and
             the SV40 early polyA sequence into pKSI+. The XhoI-SalI fragment
             of pUHC13-3 (contg. 7 copies of the tet operator upstream of minimal
             promoter Tetp) was cloned upstream of the splice/polyA sequence of
             pSplice-PATo to form preT-splice. A modified tetracycline
             transactivator (tTrAk) gene was cloned into preT-splice to form Pterr-
             tTrAk. The construct provides rapid, reversible control of gene
             expression in eukaryotic cells or transgenic animals, e.g. for
             heterologous protein, RNA or antisense sequence production.
```

5178 BP; 138; A; 1169 G

Sequence	5178 BP;	1383 A;	1169 C;	1240 G;	1386 T;
1	ATG	ATG	ATG	ATG	ATG
2	ATG	ATG	ATG	ATG	ATG
3	ATG	ATG	ATG	ATG	ATG
4	ATG	ATG	ATG	ATG	ATG
5	ATG	ATG	ATG	ATG	ATG
6	ATG	ATG	ATG	ATG	ATG
7	ATG	ATG	ATG	ATG	ATG
8	ATG	ATG	ATG	ATG	ATG
9	ATG	ATG	ATG	ATG	ATG
10	ATG	ATG	ATG	ATG	ATG
11	ATG	ATG	ATG	ATG	ATG
12	ATG	ATG	ATG	ATG	ATG
13	ATG	ATG	ATG	ATG	ATG
14	ATG	ATG	ATG	ATG	ATG
15	ATG	ATG	ATG	ATG	ATG
16	ATG	ATG	ATG	ATG	ATG
17	ATG	ATG	ATG	ATG	ATG
18	ATG	ATG	ATG	ATG	ATG
19	ATG	ATG	ATG	ATG	ATG
20	ATG	ATG	ATG	ATG	ATG
21	ATG	ATG	ATG	ATG	ATG
22	ATG	ATG	ATG	ATG	ATG
23	ATG	ATG	ATG	ATG	ATG
24	ATG	ATG	ATG	ATG	ATG
25	ATG	ATG	ATG	ATG	ATG
26	ATG	ATG	ATG	ATG	ATG
27	ATG	ATG	ATG	ATG	ATG
28	ATG	ATG	ATG	ATG	ATG
29	ATG	ATG	ATG	ATG	ATG
30	ATG	ATG	ATG	ATG	ATG
31	ATG	ATG	ATG	ATG	ATG
32	ATG	ATG	ATG	ATG	ATG
33	ATG	ATG	ATG	ATG	ATG
34	ATG	ATG	ATG	ATG	ATG
35	ATG	ATG	ATG	ATG	ATG
36	ATG	ATG	ATG	ATG	ATG
37	ATG	ATG	ATG	ATG	ATG
38	ATG	ATG	ATG	ATG	ATG
39	ATG	ATG	ATG	ATG	ATG
40	ATG	ATG	ATG	ATG	ATG
41	ATG	ATG	ATG	ATG	ATG
42	ATG	ATG	ATG	ATG	ATG
43	ATG	ATG	ATG	ATG	ATG
44	ATG	ATG	ATG	ATG	ATG
45	ATG	ATG	ATG	ATG	ATG
46	ATG	ATG	ATG	ATG	ATG
47	ATG	ATG	ATG	ATG	ATG
48	ATG	ATG	ATG	ATG	ATG
49	ATG	ATG	ATG	ATG	ATG
50	ATG	ATG	ATG	ATG	ATG
51	ATG	ATG	ATG	ATG	ATG
52	ATG	ATG	ATG	ATG	ATG
53	ATG	ATG	ATG	ATG	ATG
54	ATG	ATG	ATG	ATG	ATG
55	ATG	ATG	ATG	ATG	ATG
56	ATG	ATG	ATG	ATG	ATG
57	ATG	ATG	ATG	ATG	ATG
58	ATG	ATG	ATG	ATG	ATG
59	ATG	ATG	ATG	ATG	ATG
60	ATG	ATG	ATG	ATG	ATG
61	ATG	ATG	ATG	ATG	ATG
62	ATG	ATG	ATG	ATG	ATG
63	ATG	ATG	ATG	ATG	ATG
64	ATG	ATG	ATG	ATG	ATG
65	ATG	ATG	ATG	ATG	ATG
66	ATG	ATG	ATG	ATG	ATG
67	ATG	ATG	ATG	ATG	ATG
68	ATG	ATG	ATG	ATG	ATG
69	ATG	ATG	ATG	ATG	ATG
70	ATG	ATG	ATG	ATG	ATG
71	ATG	ATG	ATG	ATG	ATG
72	ATG	ATG	ATG	ATG	ATG
73	ATG	ATG	ATG	ATG	ATG

Query Match 43.08; Score 2222; DB 27; Length 5178;

Query Match	43.08;	Score 2222;	DB 27;	Length 5178;
-------------	--------	-------------	--------	--------------

Best Local Similarity	100.0%	Pred	No. 0 mis+0%
Matches	222	Conservative	0, Mismatches 0, Indels 0, Gaps 0;

2957	Db	ctcgaaggaggagcccggtaccacagcttttgctcccttaagtaggggttaattgcgcgctt	3016
2958	QY	tt	3017
2959	QY	ctcgaaggaggagcccggtaccacagcttttgctcccttaagtaggggttaattgcgcgctt	3018
3017	Db	ggcgtaacataggtcatagctgttctcgtgtgaaattgttatccactcacaaattccaca	3076
3018	QY	tt	3077
3019	QY	ggcgtaacataggtcatagctgttctcgtgtgaaattgttatccactcacaaattccaca	3078
3077	Db	caaatatcagagcgggaagcataaaggtgtaaagcctggagtgccataatcgaagtagagtaact	3136
3078	QY	tt	3137
3079	QY	caaatatcagagcgggaagcataaaggtgtaaagcctggagtgccataatcgaagtagagtaact	3138
3137	Db	cacattaaattgcggttgcgctcactgaccccgctttccactcaggaataaacctgtcgtccagct	3196
3138	QY	tt	3197
3139	QY	cacattaaattgcggttgcgctcactgaccccgctttccactcaggaataaacctgtcgtccagct	3198
3197	Db	gcattaaataatcgcccaacgcgcgggagagcgcggtttgcgtatggggcgacctccgc	3256
3198	QY	tt	3257
3199	QY	gcattaaataatcgcccaacgcgcgggagagcgcggtttgcgtatggggcgacctccgc	3258
3257	Db	tccctgcgttcgctgctcgcttgccgtgcggctgcggctgcggcgagcggtatcagctca	3316
3258	QY	tt	3317
3259	QY	tccctgcgttcgctgctcgcttgccgtgcggctgcggctgcggcgagcggtatcagctca	3318
3317	Db	ctcaaggcggttaattacgctttatccaggaatcagggatataacgcaggaataaacatgtg	3376
3318	QY	tt	3377
3319	QY	ctcaaggcggttaattacgctttatccaggaatcagggatataacgcaggaataaacatgtg	3378
3377	Db	gcataaagcgcgagcaaaagccaggaacgataaagcgccgcttgtcggcgcttttcca	3436
3378	QY	tt	3437
3379	QY	gcataaagcgcgagcaaaagccaggaacgataaagcgccgcttgtcggcgcttttcca	3438
3437	Db	taggctcgccccctgcgaagcatcacaaatctcagcgtcaagtcagaggtggcgaaa	3496
3438	QY	tt	3497
3439	QY	taggctcgccccctgcgaagcatcacaaatctcagcgtcaagtcagaggtggcgaaa	3498
3497	Db	ccgcagagactataaagataccagcggtttccccctggagctccctcgtcgcgctccc	3556
3498	QY	tt	3557
3499	QY	ccgcagagactataaagataccagcggtttccccctggagctccctcgtcgcgctccc	3558
3557	Db	tattccgacctgcgcttaccggataaccttcgcgctttctcccttcgggaagcgctgc	3616
3558	QY	tt	3617
3559	QY	tattccgacctgcgcttaccggataaccttcgcgctttctcccttcgggaagcgctgc	3618
3617	Db	gctttctcgaagctcacgctgtaggatctcaggttcaggttaggtgattcgctcccaagct	3676
3618	QY	tt	3677
3619	QY	gctttctcgaagctcacgctgtaggatctcaggttcaggttaggtgattcgctcccaagct	3678
3677	Db	ggcgctgtgtgcaggaaccccccggttcagccccagcgctgcgccttatccggtaactatcg	3736
3678	QY	tt	3737
3679	QY	ggcgctgtgtgtgcaggaaccccccggttcagccccagcgctgcgccttatccggtaactatcg	3738
3737	Db	tcttgaagtceraacccggttaadacacagcttatccaccactgcgcagcagccactgatacag	3796
3738	QY	tt	3797
3739	QY	tcttgaagtceraacccggttaadacacagcttatccaccactgcgcagcagccactgatacag	3798
3797	Db	gattacagcagcgaggtatgcagggggtgcataaggttgtgaagtggtggcgcttaacta	3856
3798	QY	tt	3857
3799	QY	gattacagcagcgaggtatgcagggggtgcataaggttgtgaagtggtggcgcttaacta	3858
3857	Db	cagctacacataaaggacgatttgggttatctgcgctctgactgaagccacttaccctcgg	3916
3858	QY	tt	3917
3859	QY	cagctacacataaaggacgatttgggttatctgcgctctgactgaagccacttaccctcgg	3918
3917	Db	aaaaaagttgtagctcttgcagcgcaaaacacacccgcttgtagcggtgattttt	3976


```

UY 4272 AATATATACGGAGGAGGTTTACATCTGCGCCAGTCTCAATGATACAGCGAGACCC 4331
11 4634 acactcaccggtccagatcttatcaacataaaccacacacccggaagggccagagcag 6092
UY 4332 AGGCTTACAGGCTCAGATTTATATATATATATATATATATATATATATATATATAT 4391
10 4693 aatagctcctcaaatctatccggtccatccagctctatctatctgcccgaagaagctag 6752
UY 4392 AACTGCTCTGCACTTTATATATATATATATATATATATATATATATATATATATAT 4451
10 4753 aataaataatccagcttaataatcttcgcgaacgtctatccatctgctacagacatcgt 6812
UY 4452 AATAAATATTCGCGAGTTAATATCTTCCGCAACGTTCTCCCATCTGACAGCATCGT 4511
10 4872 aatctcacctctatctcttatctatctccatctctatctctctctctctctctctct 6872
UY 4512 GATCTCAGCTCTGCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 4571
10 4932 aatctacatctatccctcatctctctctctctctctctctctctctctctctctctct 6932
UY 4572 AGTATCATGATCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4631
10 4992 tctctacatctatccctcatctctctctctctctctctctctctctctctctctctct 6992
UY 4632 TGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4691
10 7052 tctctacatctatccctcatctctctctctctctctctctctctctctctctctctct 7052
UY 4692 TTTTATCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4751
10 7112 attcttaataatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 7112
UY 4752 ATTCTGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4811
10 7173 taccgagacatagacagatcttataaagctctctctctctctctctctctctctctct 7172
UY 4812 TACCGGCGGCACATACAGCAATTTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 4871
10 7232 aaaaatctcaagatctctacagctctctgctgctgctgctgctgctgctgctgctgct 7232
UY 4872 AAAACTCTCAAGGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4931
10 7292 caatctgctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 7292
UY 4932 CAATCTGATCTTCCGATCTCTTACTTTCCAGCAGGCTTCTGCTGCTGCTGCTGCTG 4991
10 7352 accaaatcccgcaaaaagcaataaagcgaacagcaagcaagcaagcaagcaagcaagca 7352
UY 4992 GCATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5051
10 7412 cctctctctctctctctctctctctctctctctctctctctctctctctctctctct 7412
UY 5052 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5111
10 7472 taatctgctctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 7472
UY 5112 TGAATGCTATTTAGAAAAATAAACAATATAGGCTTCTGCTGCTGCTGCTGCTGCTG 5171
10 7473 ac 7474
UY 5172 AC 5174

```

RESULT 3

ID T49877 standard; DNA: 6206 BP.

AC T49877

DT 28-MAR-1997 (first entry)

DE Autoclavable vector plasmid pTet-tTAk.

KW pTet-tTAK: pTet-Splice, tetracycline transactivator, promoter;

KW gene expression; transgenic animal; animal model; drug screening.

KW vector; ds.

OS Synthetic.

```

PH Key Location/Qualifiers
FT cds complement (2466..3476)
  /*tag= a
  WO9640946-A1.
  PN 19-DEC-1996.
  PF 07-JUN-1996; U10109.
  PR 07-JUN-1995; US-474169.
  PA (UYA) UNIV YALE.
  PI Schatz DG.
  DR WPI; 97-077273/07.
  DR P-PSDB; W08474.
  PT Nucleic acid encoding tetracycline transactivator fusion protein
  PT provides rapid and reversible control of gene expression, e.g., for
  PT creating animal models for drug screening
  PS Example 1: Fig 10b-g; 82pp; English.
  CC Autoclavable plasmid pTet-tTAK (T49877) contains a modified
  CC tetracycline transactivator gene, tTAK, under control of minimal
  CC promoter, Tetp. It was obtd. by cloning the tTAK gene into vector
  CC pTet-Splice (T49876). Tetracycline prevents tTA from binding to
  CC Tetp, preventing expression of tTA and luciferase reporter gene.
  CC When tetracycline is removed, tiny amounts of tTA protein (which
  CC may result from leakiness of Tetp) bind to tet-op and stimulate
  CC expression of the tTAK gene. A positive feedforward loop is
  CC initiated leading to higher levels of tTA and luciferase. The
  CC autoregulatory control system allows rapid, reversible control of
  CC gene expression in eukaryotic cells or transgenic animals, e.g., for
  CC heterologous protein, RNA or antisense sequence production, or for
  CC creating animal models of human disease.
  SQ Sequence 6206 BP; 1620 A; 1443 C; 1486 G; 1657 T;

```

```

Query Match 42.9%; Score 2220; DB 27; Length 6206;
Best Local Similarity 100.0%; Pred. No. 0,000,000;
Matches 2221; Conservative 0; Mismatches 1; Indels 0; Gaps
DB 3985 ctcaagggggggccgctacccacattttctctctctctctctctctctctctctct 4044
UY 2952 cttcgaacacacacacacacacacacacacacacacacacacacacacacacacac 4011
DB 4045 gacataatcatgctcatgctctctctctctctctctctctctctctctctctctct 4104
UY 3012 gacgtaatactatgctctctctctctctctctctctctctctctctctctctct 4071
DB 4105 caacatacagcccggaagcacaataatctaaagctctctctctctctctctctctctct 4164
UY 3072 caacatacagcccggaagcacaataatctaaagctctctctctctctctctctctct 4131
DB 4165 cactatctctctctctctctctctctctctctctctctctctctctctctctctct 4224
UY 3132 cactatctctctctctctctctctctctctctctctctctctctctctctctct 4191
DB 4255 gacataatgactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 4284
UY 3152 gactatctctctctctctctctctctctctctctctctctctctctctctctct 4251
DB 4285 tctctctctctctctctctctctctctctctctctctctctctctctctctctct 4344
UY 3252 tttctctctctctctctctctctctctctctctctctctctctctctctctct 4311
DB 4345 ctcaagggcggttaatacctctctctctctctctctctctctctctctctctctct 4404
UY 3312 ctcaagggcggttaatacctctctctctctctctctctctctctctctctctctct 4371
DB 4405 aqcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaa 4464
UY 3372 agcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaaagccagca 4431
DB 4465 taggctcggcccccctgacagacatcacaaaaatcacagctctctctctctctctct 4524
UY 3432 taggctcggcccccctgacagacatcacaaaaatcacagctctctctctctctct 4491
DB 4525 ccccaacagactataaagactacacacattctctctctctctctctctctctctct 4584
UY 3492 cccgacagagactataaagactacacacattctctctctctctctctctctctct 4551

```

Db 4585 ttttccgaacccctgcgcttaccgataccctgccccttctcccttcggaagcgtgac 4644
QY 3552 TGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGGCTTCTCCCTTCGGNAGCGTGGC 3611
Db 4645 gcttctcctacgtcagctgtaggtatctcagttctggtgtaggtctgctccaaagt 4704
QY 3612 GCTTTCTTATAGTCTACGCTGTATGATGATGATGATGATGATGATGATGATGAT 3671
Db 4705 ggcctgtgtacagaaacccccccttcagcccaaccccttcagcttatccggttaactatcg 4764
QY 3672 GGGCTGTGTGACAAACCCCGCTTACGCTGTGCTTATCGGTAATATCG 3731
Db 4765 tcttgatcccaaccccgtaagacacagcttatccgcaactggcagcagccactggtaacag 4824
QY 3732 TCTTGAGTCCACCCGTAAGACACGACTTATCGCCACTGGCAGCCACTGGTAACAG 3791
Db 4825 gattaacagagcagagtatgagccggtgctacagagttcttgaagtggtgagcetaacta 4884
QY 3792 GATTAGCAGAGCGAGGATATGAGGCGGTGCTACAGAGTTCITSAAGTGGTGGCTTAAGTA 3851
Db 4885 cggctacactagaagacagctatttggtatctgcgctctgctggaagccagttacacctcgg 4944
QY 3852 CGGCTACACTAGAGACAGATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGG 3911
Db 4945 aaaaagagttggttagcttctgctccgcaacaaacccacccgctgtagcggtggttttt 5004
QY 3912 AAAAAGAGTTGTTAGCTCTTGATCCGGCAACAAACCCACCGCTGGTAGCGTGGTTTTT 3971
Db 5005 tgtttgcaacagcagatatacgccgagagaaaaaagatctcgaagatcccttcaatt 5064
QY 3972 TGTTTGCAAGCAGCAGATTACGGCGGAGAAAAAAGATCTCAAGAGATCCCTTTGATCTT 4031
Db 5065 ttctacgggtctgacgctcagtggaacgaaactcaactgaaggtatttggtcag 5124
QY 4032 TTCTACGGGTCTGACGCTCAGTGGACGAAACTCACGTTAAGGATTTTGGTCTAGAG 4091
Db 5125 attataaaagatcttcacctagatcccttttaaaattaaaatgaagtttaaatcaat 5184
QY 4092 AITATCAAAAGAGATCTTCACTAGATGCTTTTAAATTAATAATGAAGTTTAAATCAAT 4151
Db 5185 ctaagtatataagtaaaacttggtctacagttaccacttaacttaactgaaggaacc 5244
QY 4152 CTAAGTATATAGTAATACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGACCC 4211
Db 5245 tatctcagcgtctgctatttctgctcctcctcagttgctcactcccgctgctagat 5304
QY 4212 TATCTCAGCGATCTGCTATTTCGTTCACTCCTAGTTGCTGACCTCCGCTGCTAGAT 4271
Db 5305 aactacgatacgggagggcttacatctggcccaagctgctgaatgataccgagagacc 5364
QY 4272 AACTACGATACGGGAGGGCTTACCATCTGGCCCACTGCTGCAATGATACCGCGAGAGCC 4331
Db 5365 acgctcaccggtctcagattatacgaataaaacagcagccggaagccgagcgcag 5424
QY 4332 ACGTCTACCGGCTCCAGATTATCAGCAATAAACACAGCAGCGCGGAGGCCGAGCGGAG 4391
Db 5425 aagtgctcctcgaactttatccgctccatccagctctataattgttgcgggaagctag 5484
QY 4392 AAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTGTTGCGGGAAGCTAG 4451
Db 5485 aataagtagtcccaatttaattgttggcaacagctgttgcattgctacaggaatcgt 5544
QY 4452 AGTAAGTAGTTTCGCACTTAATAGTTGGGCAACGTTGTGGCAATGCTACAGGATCGT 4511
Db 5545 ggtgtcagctcgtcgtttggttgggtcttcattcagctccggttcccaacagatcaagcg 5604
QY 4512 GGTGTACGCTCGCTGCTTGGTATGCTTCATTCAGCTCGGTTCCCAACGATCAAGCGG 4571
Db 5605 agttacatgatccccatgttgtgcaaaaaagcggttagctccttccctccagatcgt 5664
QY 4572 AGTTACATGATCCGCAATGTTGTGCAAAAAAGCGGTTAGTCTCTCTGCTCCGCTCGATCGT 4631

RESULT 4

ID Q40279 standard: DNA: 4145 BP.
AC Q40279;
DT 02-AUG-1993 (first entry)
DE Sequence of clone pS2gpt-S4.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..2225
FT /tag= a
FT /label= pN2gpt-S4
FT /note= "Posn. 1 corresp. to the first nucleotide
G '5'-TGGCACTTT TCGGGGAAAT-3'."
FT misc_feature 2227..2236
FT /tag= b
FT /label= SmaI adaptor
FT 2396..2851
FT /tag= c
FT /label= E. coli gpt gene
FT complement (3081..3323)
FT /tag= d
FT /label= vaccinia p7.5 promoter
FT 3358..3451
FT /tag= e
FT /label= S4 of pN2gpt-S4
FT /note= "Oligo P-aptP(9) "
FT misc_feature 2237..4145
FT /tag= f
FT /label= pN2gpt-S4
FT AU9221269-A.
PN 04-MAR-1993.
PD 25-AUG-1992. 921269.
PF

PR 26-AUG-1991: US-750080.
PB 20-JUN-1992: US-914738.
PA (IMMO) IMMUNO AG.
PI Dörner F., Falkner FG, Pfeleiderer M, Scheiflinger F.
DR WPI: 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
FT direct molecular cloning of modified RNA molecule contg.
PI cytoplasmic DNA virus genome
PS Example: Pages 160-162; 206pp; English.
CC Plasmids pN2apt-S3A and pN2apt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gptA and pN2-gptB which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. The plasmid
CC pN2apt-S4 was digested with XbaI and ligated with a SmaI-adaptor
CC inactivating the XbaI and creating a SmaI site. The resulting
CC plasmid was designated pS2apt-S4. The S4-promoter segment of
CC plasmid pS2apt-S4 was removed by cleavage with PstI and HpaI and
CC replaced with a 172 bp pSti-HpaI P2-promoter segment. The final
CC plasmid was designated pS2apt-P2.
SQ Sequence 4145 BP; 1067 A; 1007 C; 1039 G; 1032 T;
Query Match 42.6%; Score 2202; DB 7; Length 4145;
Best Local Similarity 100.0%; Prod No. 0.00e+00;
Matches 2202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 atggcaacttttcgggaaatgtgcgcgaacccctatttttttcttaataacatt 60
Cp 5173 GTGAGACTTTTCGGGCAAAATGAGGGAACCTATTGTGTTATTTTCTAAATACATT 5114
Db 61 caaatatgtatccgtcatgaacaataaacctgatgaataatgcttcaataattgaaaa 120
Cp 5113 CAAATATGATATCCGCTCATGACACATAAACCTGATAAATGCTCAATATATTGAAAA 5054
Db 121 gaaagaatgaatgaattcaacatttcccgctgcgccttattcccttttttgcgcatttt 180
Cp 5053 GGAAGAGTATGATGATTAAATTTTCCTGCTGCGGCTTATTCCTTTTTCGGGCATTTT 4994
Db 181 gcttctcttttttgcaccagaaacgcctgggtgaaagttaaagatgctgaagatcagt 240
Cp 4943 GCTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4934
Db 241 taattgcacagatgggtttacatgcgaactggatctcaacagcggtaagatccttgagatt 300
Cp 4933 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4874
Db 301 ttccqcccccgaadaacgcttttccaatgatgagcacttttaaaattctctatgtgcgcg 360
Cp 4873 TTGCGCGCGAAGAAAGCTTTTCTAAATGATGACACTTTTAAAGTCTCTGCTATGTGCGCGG 4814
Db 361 tattatcccgatttgaacgcgggaagagcaactcggtcgccgcatacacactattctcaga 420
Cp 4813 TATTATCCGCTATTGACGCGCGCGAAGGACCACTCGGCTGCGGCTACACTATTCTCAGA 4754
Db 421 atgaacttgdttaagtactaccagctcacagaaaaacatcttaccggatggcatgacatga 480
Cp 4753 ATGACTTGCTTGAAGTACTCATATGCTACAGAAAAAGCATCTTACGGATGGCATGACATA 4694
Db 481 aaaaattatgaactgaccataaccatgaatgataaacactgcggccaaacttaacttctga 540
Cp 4693 GAATAATTGAAATGAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 4634
Db 541 caacqatcgaagagcgaagagcgaactaacgcgtttttttcacacaacatggggatcatgaa 600
Cp 4633 CAAAGATGCGAG 4574
Db 601 ctgccttgatcgttgggaacgcggactgaatgaagccataccaaacagcagcgctgaaca 660
Cp 4573 CTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4514
Db 661 ctacgatgctatgaactggcaaacgcttgcgcgaactatttaactggcggaactactta 720
Cp 4513 CCACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4454

Db 721 ctctagcttcccgccgaacaataatgaactgagatgagagagagagagagagagagagag 780
Cp 4453 CTCTAGCTTCCCGGCAACAATTAATAGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 4394
Db 781 tctgcgcctgcgccttcccgctgcgccttcttctgaataaactctgcgcgcgcgcgcgcgc 840
Cp 4393 TTCTGCGCTCGGCGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4334
Db 841 atgggtctcgcggtatcattgcacgcctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900
Cp 4333 GTGCGCTCTCGGCTATTCATTGCAATATGCGGCAACAATGGAAGAGAGAGAGAGAGAG 4274
Db 901 ttatctac 960
Cp 4273 TTATCTAC 4214
Db 961 taggtgcctcactgattaaac 1020
Cp 4213 TAGGTGCTCTACTGATTAAAGCATTTGGTAACTGTGCAACACACACACACACACAC 4154
Db 1021 agattgatttaaaacttcttatttaaaadgaatctaggtgaadgaatcttcttcttctt 1080
Cp 4153 AGATTGATTAAAACTTCATTTTAAATTAAGAGGATCTAGCTCAACACATCTCTTCTATA 4094
Db 1081 atctcatgac 1140
Cp 4093 ATCTCATGACCAAAATCTTAACTGAGTTCAGTTTTCGTTCCACTGAGGTGACAGACAC 4034
Db 1141 aaagatcaaaaggaatcttcttgaatccttcttctgcgcgcgcgcgcgcgcgcgcgcgc 1200
Cp 4033 AAAGATCAAAAGGATCTCTTGAATGCTTTTTCGCGGCTAACTGCTGCTGCTGCTGCT 4974
Db 1201 caaaaaaac 1260
Cp 3973 CAAAAAAAC 4914
Db 1261 ttcgaagatgaactggcttcagcaaacacacacacacacacacacacacacacacacac 1320
Cp 3913 TTCTGCAAGGTAACTTGGGTTCTGAGCAGACGCGGAGATACCAAACTACTGTCTTCTAG 4854
Db 1321 cgtagtttaggcac 1380
Cp 3853 CGTACTTAGGCGACACACTTCAAGAACTCTCTAGCAGCGGCTACATACCTGCTCTCTAA 4794
Db 1381 tctcttaccagtgagctgcctccagtagagataagctgctgttcttaccggttgaatcga 1440
Cp 3793 TCCTGTTACCGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4734
Db 1441 gacgatagttaccggataaagcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1500
Cp 3733 GACGATAGTTACCGGATAAGGCGTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4674
Db 1501 ccagtttggaggaac 1560
Cp 3673 CCACGCTTGGAGCGAAGCAGCTACACGCAACTGACATACCTTACAGCGCTGAGCTATGAGA 4614
Db 1561 gcgcacacgttcccgaaaggaagaaacacacacacacacacacacacacacacacacac 1620
Cp 3613 GCGCGCAGCGCTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4554
Db 1621 caagagagcgcacgagagagccttccaaaggaagaaacacacacacacacacacacac 1680
Cp 3553 CAGGAGAGGCTGCTGCAAG 4494
Db 1681 gtttcgcac 1740
Cp 3493 GATTTTGGCAGCTCTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4334
Db 1741 tatgaaaaaac 1800
Cp 3433 TATGAAAAAACGCGACGCAACCGCGCGCTTTTACGGTTCCTGGGCTTTTCTGCTGCT 4374

PI Control of genes in transgenic plants - using an upstream activating
 PT sequence activated by a transactivating protein expressed using a
 separate promoter

PS Example: Fig 5F: 48bp, English.
 CC Reporter plasmid pUMIGIT (74347), of pUAS Minimal promoter Gus
 CC Int Terminator, contains a beta-glucuronidase (GUS) reporter
 CC gene under the control of the 46S cauliflower-mosaic virus minimal
 CC promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream
 CC activating sequence (UAS) of pUMIGIT is activatable by yeast
 CC transactivating protein GAL4. In a novel method for controlling
 CC gene expression, a first transgenic plant carrying a gene encoding
 CC a desired phenotype (herbicide resistance of polyhydroxybutyrate
 CC prodn.) operatively linked to a UAS recognition site is pollinated
 CC by a second transgenic plant carrying Gal4 DNA (see also 743136)
 CC The transgene is fully expressed in F1 hybrid plants but segregates
 CC apart in subsequent generations improving environmental safety
 CC Sequence 5544 bp: 1400 A: 1355 G: 1387 G: 1363 T:

Query Match 42 % Score 2202; DB 25; Length 5534;
 Best local Similarity 100.0%; Prod. No. 0.00e+00;
 Matches 2202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3329 atdcaacttttgggaataatgacgaacccctatttttttataatacat 3388
 Cp 5173 GTGGACATTTTGGGAAATGTGGGGAACCCCTATTGTGTTTATTTCTTAAATACATT 5114
 Db 4389 caaatatcatccatcaatcaacataacccctataataatcttcaataataatgaaaaa 3448
 Cp 5113 CAATATGATATGGCTATGAGAAATATATACCTGGATGAATAGCTTCAATATATGAAAAA 5054
 Db 4449 gaaagatataatattcaacatttccctgtccctgtccctgtttttttggcgcatttt 3508
 Cp 5053 GAAGACATATGATTAAATTTCCCTGTCCCTGTCCCTGTATTCCTTTTTCGSCATTTT 4994
 Db 4509 gcttctatttttgcacccagaaagactgtgtaaaatgaaatgctctgaaatcaagt 3568
 Cp 4994 GCTTCTGCTTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4934
 Db 4569 tgaatcacatgggttacatcgaactggaatctcaacagcggtaagatccctcaagatt 3628
 Cp 4934 TGAGTCAGACATGGGTTTACATCGAATGCGATCTCAATAGCGGTAAAGATCCTTTGAGAGTT 4874
 Db 4629 ttcccccgaagaagatttttcaatgatgagcaatttttcaagttctgctatgttgagcgg 3688
 Cp 4873 TTCCGGCGGAAGAAAGCTTTTTCATGATGAGCACTTTTAAATTTTGTATGSGGCGGG 4814
 Db 3689 tatatcccatattacacggcgagcagagcaactgggtcgagcgtatataactattctcaga 3748
 Cp 4814 TATTATCGGATATTAAGTGTGGGTAAGAGCAATGGGTCGGGCAATACATATTTCTCAGA 4754
 Db 3749 atgatttggttgagtagtctcaggatcagagaagaagcatcttacgcatggcatgacatgaa 3808
 Cp 4753 ATCACTTGATTCAGTACTTACAGTCAACAGAAAGCATCTTACGCGATGGCATGACAGTAA 4694
 Db 4809 gaaatatacatgctgcataaactatgagtgatataacactgcggccaacttactctga 3868
 Cp 4694 GAGAAATATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4634
 Db 4869 caacgatccgaggaacgaagagctaacccgtttttttcaacaacatagggaatcatataa 3928
 Cp 4634 CAAGCATGTCAGATTAAGCAATTAATATATATATATATATATATATATATATATATATAT 4574
 Db 4929 ctgagtttgatgttggaacggaggttgatgagagcctatacccaacgacgagtgtaaca 3988
 Cp 4574 CTGAGTTTATGCTTCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTA 4514
 Db 4989 ccacgatccctatagaaatggaaacagcttgcggaactattaactggcgaactactta 4048
 Cp 4513 CCAATGATGCTTCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAATCTGTAAT 4454
 Db 4049 ctctagcttccgggaacgaatgaatgaactggaatgagagcagataaagtgcaggaacac 4108
 Cp 4453 CTCTAGCTTCTGCGGCAACCAATTAATAGATGATGCAATGCAATGCAATGCAATGCAATGCA 4394

Db 4100 ttctgagctcagcgccttcagcgtgagctgagctgagctgagctgagctgagctgagctgagc 4168
 Cp 4393 TTCTGCGCTGCGGCGCTTCCGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 4344
 Db 4169 atggggtctcgcggtatgagctgagcagcagcagcagcagcagcagcagcagcagcagcagc 4228
 Cp 4344 GTGGATATCGGCGGTATATATGCAAGACTGGGCGGTAGAGGAGGCTGGGCTGGGCTGGGCT 4274
 Db 4229 ttatctacacgacgaggaatccgagcagctatgataaagaaatataaagatcagcagaa 4288
 Cp 4273 TTATTTATAAGAGAGGGAAT 4214
 Db 4289 taggtgcctcactaatataagcattgagcagcagcagcagcagcagcagcagcagcagcagc 4348
 Cp 4213 TAATGCTGCTCAATGATTAAGGATTTGATTAAGGATTTGATTAAGGATTTGATTAAGGAT 4314
 Db 4349 agattgatttaaaacttcatctttaaatttaaaagagctgagcagcagcagcagcagcagcagc 4408
 Cp 4153 AGATTGATTAAAGCTTCAATTTTAAATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTA 4364
 Db 4409 atctcatgaccaaatacccttaacagcagcagcagcagcagcagcagcagcagcagcagcagc 4468
 Cp 4093 ATCTCATGACCAAAATCCCTTAACGCTGAGTTTGCTTCAATGAGGCTCAATGAGGCTCA 4384
 Db 4469 aaaaatcaaaagcattctcttgaatcctttttctccagcagcagcagcagcagcagcagcagc 4528
 Cp 4033 AAAATATTAAGAGATTTTCTGAGATGCTTTTCTGAGGTAATTTCTGAGGTAATTTCTGAG 4574
 Db 4529 caaaaaaccacgcgctacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4588
 Cp 3973 CAAAAAACCAGGCTAAGAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4514
 Db 4589 ttccgaaggttaacttggttcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4648
 Cp 3913 TTCCGAAGGTAACTGGCTTCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4594
 Db 4649 ctagttaggcac 4708
 Cp 3853 CGTAGTTAGGCTCACTTCAAGAGCTCTGTAGTACCGGCTATATATATATATATATATATAT 4754
 Db 4709 tctgtattaccagtgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4768
 Cp 3793 TCTGTATTACCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4734
 Db 4769 aacgatcaattacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4828
 Cp 3733 GACGATAGTTACCGGATAAAGGCGGAGGCTGGGCTGAAAGGGGCTGGGCTGGGCTGGGCT 4674
 Db 4829 ccaagcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4888
 Cp 3673 CCAGCTTGGAGGAGAGAGAGCTTACATGCAAGCTGCAATACAGAGAGAGAGAGAGAGAG 4614
 Db 4889 ggcgcagccttcccaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4948
 Cp 3613 GGGGTAAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTA 4864
 Db 4949 caggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4968
 Cp 3553 TAGGATAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTAAGGCTTGGGTA 4914
 Db 5009 ggtttcgccacactctgagcttgagcagcagcagcagcagcagcagcagcagcagcagcagc 5068
 Cp 3493 GGTTCGCGCACCTCTGACTTCAGCGGTGAGTTTCTGATCTGCTTACGCGGCTGGAGGCT 5014
 Db 5069 tatgaaataacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5128
 Cp 3433 TATGGAAAAAGCGTAAGTAAAGGAGGCTTTTAAAGGCTGAGGCTTTTAAAGGCTTTTAA 4974
 Db 5129 ctccatattcttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5188
 Cp 3373 CTCACATGTTCTTCTGCTGGGTTATGCGGCTGATTCGCTGATGCTGCTGCTGCTGCTGCTG 4914

Db 5189 atggaactataccgtctgcccgaacacgaagacagcagcagcgagcagtgcaagtaagcaggg 5248
 Cp 3313 AGTGAGCTGATAGCGTGGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3254
 Db 5249 aagcggaaagcgcgaataacgaacacccgcctctcccgccgcttgccgagcattcaatt 5308
 Cp 3253 AAGCGGAAGAGGCGCAATACGCAACGCGCTCTCCCGCGCGCTGGCGGATTCATTAA 3194
 Db 5309 gcaactggcagcagcaggttcccgactgagaagagggcgagcagcagcagcagcagcagc 5368
 Cp 3193 GTAAATTTGTAACAGGTTTCCGACATGGAAGGCGGAGGAGGAGGAGGAGGAGGAGG 3134
 Db 5369 tgagttagctcactcattaggcaaccccgaggtttacacttttatctctccgctctgattg 5428
 Cp 3133 TCAGTTAGCTCACTCATTAGGACCCAGGCTTACACTTTATGCTTCGGGCTGATGT 3074
 Db 5429 tgtgtgaatttgagcgataacaaatttcacagaggaacacagctatgacacattcaag 5488
 Cp 3073 TGTGTGAATTGTGAGCGGATTAACAAATTCACAGAGGAGGAGGAGGAGGAGGAGGAGG 3014
 Db 5489 ccaagcgcgaatttaacctcactaaagggaacaaaagctgg 5530
 Cp 3013 CTAAAGCGCGCAATTACCTGACACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2972

RESULT 8
 ID Q13578 standard: DNA; 3681 BP.
 AC Q13578;
 DT 02-DEC-1991 (first entry)
 DE Plasmid pKSEL5.
 KW recombinant expression plasmid; phage phi X174; protein E;
 KW bacterial ghost; ss.
 OS Synthetic.
 PN WO9113155-A.
 PD 05-SEP-1991.
 PF 19-FEB-1991; E00308.
 PR 24-FEB-1990; DE-005874.
 PA (BOEF) BOEHRINGER MANNHEIM GMBH.
 PI Lubitz W, Szostak MP;
 DR WPI: 91-281471/38.
 PT Carrier recombinant protein for vaccines against HIV, etc.
 PT obtd. by expressing fusion protein gene in gram negative bacteria
 PT and gene that encodes for lytic membrane protein
 PS Example 1; page 33-34; 45pp; German.
 CC Plasmid pKSEL5 comprises three multiple cloning sites (mcs1-3), the
 CC ampicillin resistance gene, regions of the lac operon and a
 CC partial phage phi X174 E' target sequence. This plasmid can be used
 CC as a carrier vector according to the invention for the insertion of
 CC sequences encoding viral antigenic peptides, e.g. the HIV gp41
 CC sequence was isolated as a HincII/PvuII fragment from plasmid
 CC pHF14. The fragment also included a linker sequence and the last 45
 CC codons of gp120. The fragment was ligated to pKSEL5 which had been
 CC cleaved with AccI and end-filled. Alternatively, a new carrier
 CC vector, denoted pMTV1 (see p13576), can be constructed from pKSEL5
 CC which has been partially digested with SspI. It is ligated to a
 CC SspI-draII fragment from pML1 (see Q13577). Foreign sequences are
 CC cloned into mcs2.
 SQ Sequence 3681 BP; 949 A; 917 C; 932 G; 883 T;

Query Match 42.5%; Score 2201; DB 2; Length 3681;
 Best Local Similarity 99.4%; Pred. No. 0.00e+00;
 Matches 2114; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 1452 tcaacctcagggggggcccggtaccacagcttttgcctcttagtggggttaattccg 1511
 QY 2947 TCGAATCTAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3006
 Db 1512 agcttgcgtaatcagtcagtagctgttctctgtgtaattgttccgcacaaatt 1571
 QY 3007 CGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACA 3066
 Db 1572 ccacacataagacccggaagcacaagaagttaagcctgggtgcctaataagtgagg 1631

QY 3067 CCACACAACTATAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3126
 Db 1632 taactcacatttaattggtctgctctcactgctgctgctgctgctgctgctgctgctgctg 1591
 QY 3127 TAACTCACATTAAATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 3186
 Db 1692 cagctgcaatttaattggtctgctgctgctgctgctgctgctgctgctgctgctgctgct 1751
 QY 3187 CAGCTTGCATTAAATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3246
 Db 1752 tccgttctcgtcactgactgctgctgctgctgctgctgctgctgctgctgctgctgctg 1811
 QY 3247 TCCGTTCTCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3306
 Db 1812 gctcactcaaaagcggtaatacaggttatccacaatacaggggataacgcaaaaataac 1871
 QY 3307 GCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATACCGCAGGAAGAC 3366
 Db 1872 atgtgagcaaaagcggcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1931
 QY 3367 ATGTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAG 3426
 Db 1932 ttccataggtcggccccctcgcagcagcagcagcagcagcagcagcagcagcagcagcag 1991
 QY 3427 TTCCATAGGCTCCGCCCTCTCAGAGCATCACAAAAATCAGAGCTCAAGTCAAGAGTGG 3486
 Db 1992 cgaacccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2051
 QY 3487 CGAAACCCGACAGGACTATAAGATACAGAGCGTTTCCGCTGGAAGCTCCCTGCTGCTG 3546
 Db 2052 tctcctgttccgacccctccgcttaccggataacctgtccgcttcccttcccttccggaagc 2111
 QY 3547 TCTCTGTTCGGACCTCGCGCTTACCGGATACCTGTTCGCTTTCCTCTTCGCGAAGC 3606
 Db 2112 gtggcgcttctcgaatgctcagcgtgtagtattcaggttggttggttggttggttggt 2171
 QY 3607 GTGGCGCTTCTCATAGCTACGCTGTAAGTATGTAAGTATGTAAGTATGTAAGTATG 3666
 Db 2172 aagctggcgctgtgacagcaaccccgcttccgcccccgctgctgctgctgctgctgctg 2231
 QY 3667 AAGCTTGGGCTGTGTGACAGAACCCCGCTTCCGCGGACGCTGCGGCTTATCGCGTAA 3726
 Db 2232 tatctcttgagtcacaccccggtgaagacacagcattatcgccactggcagcagcagcagc 2291
 QY 3727 TATCGTCTTGTAGTCCACCCCGTAAGACACAGCATTATCGCCACTGGCAGCAGCAGCT 3786
 Db 2292 aacagattagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2351
 QY 3787 AACAGGATTAGCAGAGCAGGATATGAGCGGCTGTACAGAGATTCTTGAAGTGGTGGCT 3846
 Db 2352 aactcggctcactcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2411
 QY 3847 AACTAGCGCTACACTAGAGGACAGTATTGGTATCTCGCTTCTGCTGAAGCTAGTACC 3906
 Db 2412 ttcgaaaaagagttggtagcttctgacggcaacaaacaccccgctgctgtagcggtag 2471
 QY 3907 TTCCGAAAAAGAGTTGGTAGCTTTGTATCCGCGCAACAAACCAACCGGCTGTGTAGCG 3966
 Db 2472 tttttgtttgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2531
 QY 3967 TTTTTTTTGTGCAACGACAGCATACGGGCAAGAAAAAGGATCTCAAGAAAGATCTTGG 4026
 Db 2532 atcttctcaggggtctgacgctcagtggaacacaaactcacgttaaggaattttggtc 2591
 QY 4027 ATCTTTCTACGGGCTCTACGCTCAGTGAAGCAAAACTCACGTTAAGGCAATTTGGTC 4086
 Db 2592 atgagattatcaaaagagattctcactcagtagctctttttaaattaaaattgaattttaa 2651
 QY 4087 ATGAGATTATCAAAAGAGTCTTCACTAGTAGCTTTTAAATTAATAATGAAGTTTAAA 4146
 Db 2652 tcaatcaagctatagtagtaaaactggctgacattaccactgcttaataacagtagg 2711
 QY 4147 TCAATCTAAGTATATAGAGTAAACITGGCTGACATTACCAATGCTTAATCAGTGG 4206

Db 2712 gaacatattcagagatcgtatctatcttcttcacatatttaccctgactgccctcgctg 2771
 QY 4207 ccaactatcagagatcgtatctatcttcttcacatatttaccctgactgccctcgctg 4266
 Db 2772 tagataactcagatcag 2831
 QY 4267 tagataactcagatcag 4326
 Db 2832 gaacacagctcag 2891
 QY 4327 gacacacagctcag 4386
 Db 2892 gaacacagctcag 2951
 QY 4387 gacacacagctcag 4446
 Db 2952 gctagagatcag 3011
 QY 4447 gctagagatcag 4506
 Db 3012 atcgtggtcag 3071
 QY 4507 atcgtggtcag 4566
 Db 3072 atcgtggtcag 3131
 QY 4567 atcgtggtcag 4626
 Db 3132 atcgtggtcag 3191
 QY 4627 atcgtggtcag 4686
 Db 3192 atcgtggtcag 3251
 QY 4687 atcgtggtcag 4746
 Db 3252 atcgtggtcag 3311
 QY 4747 atcgtggtcag 4806
 Db 3312 atcgtggtcag 3371
 QY 4807 atcgtggtcag 4866
 Db 3372 atcgtggtcag 3431
 QY 4867 atcgtggtcag 4926
 Db 3432 atcgtggtcag 3491
 QY 4927 atcgtggtcag 4986
 Db 3492 atcgtggtcag 3551
 QY 4987 atcgtggtcag 5046
 Db 3552 atcgtggtcag 3611
 QY 5047 atcgtggtcag 5106
 Db 3612 atcgtggtcag 3671
 QY 5107 atcgtggtcag 5166
 Db 3672 atcgtggtcag 3731
 QY 5167 atcgtggtcag 5230

RESULT 9
 ID V14340 standard: DNA: 3699 BP.
 AC V14340;

DT 19-MAY-1998 (first entry)
 DE Plasmid pBSGFP expressing green fluorescent protein.
 KW Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish
 KW genetic engineering marker; gene therapy; plasmid; ss.
 OS Synthetic.
 PN W09742320-A1.
 PD 13-NOV-1997.
 PF 07-MAY-1997; 007625.
 PR 08-MAY-1996; US-646538.
 PA (USSH) US SEC DEPT HEALTH.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Gattaris GA, Pavlakis GN, Stauber RH, Vournakis JN;
 DR WPI: 97-558982/51.
 PT New nucleic acid encoding proteins of Aequorea victoria with
 PT increased fluorescence useful as markers for detecting cellular
 PT transformation, subcellular localisation of proteins, for assessing
 PT gene therapy, mutagenicity etc.
 PS Example -; Page 67-68; 105pp; English.
 CC This sequence is a plasmid used to express the mutated green fluorescent
 CC protein (GFP) of Aequorea victoria coding sequence of the invention. The
 CC DNA of the invention that encodes a protein that is (a) a protein that
 CC has Leu at position 65 (and optionally Thr at 168 and optionally further
 CC Cys at 66) and has cellular fluorescence at least 5 times that of
 CC wild-type GFP or (b) is a blue fluorescent protein (BFP) with His at
 CC position 67 and also at least one of Leu at 65 and Ala at 164, and has
 CC cellular fluorescence at least 5 times that of BFP (Tyr67 to His). The
 CC nucleic acids can be used as markers in genetic engineering and gene
 CC therapy. They may also be used to detect and characterise regulatory and
 CC coding sequence elements that control subcellular expression and
 CC targeting of proteins. Typical applications are monitoring targeting and
 CC transport of proteins in cells; assessment of gene therapy procedures; in
 CC diagnosis (when expressed under control of a promoter induced by a
 CC particular analyte); assessment of mutagenicity of compounds; and for
 CC drug screening (where expression is controlled by the promoter of a
 CC target gene), particularly for antiviral or antiparasitic agents.
 SQ Sequence 3699 BP; 942 A, 884 C, 898 G, 975 T;

Query Match 42.5%; Score 2198; DB 39; Length 4649;
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;
 Matches 2200; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 1495 ccagctttgttccctttagtgagggtaattccagagctgagctgagctgagctgag 1554
 QY 2572 ccagctttgttccctttagtgagggtaattccagagctgagctgagctgagctgag 4041
 Db 1555 tgttccctgtgaaattgttaccctcaccacattccacacacatccagacacacacac 1614
 QY 3032 tgttccctgtgaaattgttaccctcaccacattccacacacatccagacacacac 4091
 Db 1615 taaagtgtaaagcctggggtgagctgagctgagctgagctgagctgagctgagctgag 1674
 QY 3092 taaagtgtaaagcctggggtgagctgagctgagctgagctgagctgagctgagctgag 4161
 Db 1675 cagtgcgctgttccgttggggagagctgagctgagctgagctgagctgagctgagctgag 1744
 QY 3152 cagtgcgctgttccgttggggagagctgagctgagctgagctgagctgagctgagctgag 4211
 Db 1735 gcgcggggagagcgggttggctgattggagcgtcttccgcttccgcttccgcttccgct 1794
 QY 3212 gcgcggggagagcgggttggctgattggagcgtcttccgcttccgcttccgcttccgct 4271
 Db 1795 tgcctcgttgccttgcctgagcgggtatccagctcactcacaagcagctgagctgagct 1854
 QY 3272 tgcctcgttgccttgcctgagcgggtatccagctcactcacaagcagctgagctgagct 4331
 Db 1855 tatccacagatcaggagataacagcaggaacacatgatgacacaaagccagacacagag 1914
 QY 3332 tatccacagatcaggagataacagcaggaacacatgatgacacaaagccagacacagag 4391
 Db 1915 ccaggacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1974
 QY 3392 ccaggacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 4451

PR 20-JUL-1992; US-914738.

PA (INV.) IMMUNO AG.

PI Dörner F, Falkner FG, Pfeleiderer M, Scheifflinger F;

DR WPI: 93-126461/16.

PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves

PI direct molecular cloning of modified DNA molecule contg.

PT cytoplasmic DNA virus genome

PS Example: Pages 162-164; 206pp; English.

CC Plasmids pN2qpt-S3A and pN2qpt-S4 comprise expression cassettes

CC with a selective marker. These plasmids were constructed by first

CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt

CC gene driven by the vaccinia virus P7.5 promoter, flanked by

CC several unique restriction sites including NotI. The plasmid

CC pN2qpt-S4 was digested with XbaI and ligated with a SmaI-adaptor

CC inactivating the XbaI and creating a SmaI site. The resulting

CC plasmid was designated pS2qpt-S4. The S4-promoter segment of

CC lamsid pS2qpt-S4 was removed by cleavage with PstI and HpaI and

CC replaced with a 172 bp PstI-HpaI P2-promoter segment. The final

CC plasmid was designated pS2qpt-P2.

SQ Sequence 4277 BP; 1133 A; 1021 C; 1061 G; 1060 T;

Query Match 42.5%; Score 2198; DB 7; Length 4277;

Best Local Similarity 99.9%; Pred. No. 0.00e+00;

Matches 2199; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 1 atggcacttttcggggaatgctgcggaacccctatttttttttttaataacatt 60

Cp 5173 GTGGGACATTTTCGGGAAATGTGCGGGAACCCCTATTTCGTTTCTTAATACATT 5114

Db 61 caaatatataatccgctcatgagacataaacctgaataatgctcaataatgaaaa 120

Cp 5113 CAATATATGATCCGTCATGACAAATAAGCCCTGATAAATGCTTAAATATGAAAA 5054

Db 121 ggaagatgataagatcaacattccgctatcccttattcccttttttggcgacttt 180

Cp 5053 GGAAGATGATGAGTATCAACATTTCCGTCGCTCTTATTCCTTTTTCGGGCAATTT 4994

Db 181 gcttccctgtttttgctccacgaaacgctgggtgaagtaaaagatgctgaagatcagt 240

Cp 4993 GCTTCTCTCTTTTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 4934

Db 241 gggatgacgagtaggtttacatacgaactgagatctcaacagcgtatgaagcttgaagtt 300

Cp 4933 TGGTGGGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4874

Db 301 ttccgcccgaagaaacgcttttccaatgatgagcactttttaaagtctctgctatgtggcgcg 360

Cp 4873 TTCGCTGGGAAAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4814

Db 361 tattatcccttatgacgcccgggaagagcaactcggtcgccgcatacactattctcaga 420

Cp 4813 TATTATCCCTTATTGACGCGCGGGAAGCAACTCGCTCGCGGCAATCTGCTCAGA 4754

Db 421 atgacttgatgactcaccctcaccatcagaaagcatcttaccggtggtcagatgaagtaa 480

Cp 4753 ATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 4694

Db 481 gagaattatgactgctgccataaccatgagtataacactgcggcgcaacttaactctga 540

Cp 4693 GAGAATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4634

Db 541 caacgatccggagccgaagagcttaaccgcttttttttgcacaacatgggggcatcatgaa 600

Cp 4633 CAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4574

Db 601 ctgccttgatgcttggaacacgagctgaatgaagccataccaaacgacgagcgtgaca 660

Cp 4573 CTCGCTTGATGCTTGCGGAACCGGAGCTGATGAAGCCATACCAACGACGCGTGACA 4514

Db 661 caagatgcttgatgagcaatggcaacacgcttgcgcaaacatttaactggcggaactactta 720

Cp 4513 CCACCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4454

Db 721 ctctagcttcccggaacaacttaataagatgcatgaaagcgaatgtaagaaac 780

Cp 4453 CTCTAGCTTCCCGGAACAATTAATAACAATGATGAGAGAGAGAGAGAGAGAGAG 4394

Db 781 ttctgctgcggcccttccgggtggctgtttatttctgataaaattctgaaagcgaatgaa 840

Cp 4393 TTCTGCGCTCGGCGCTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4344

Db 841 gtagctcgcgcatatcattgacgactgggacgaatgtaagacactcctcctatcctaa 900

Cp 4333 GTGGGCTTCGCGGTATCATTTGCTAGCTACTGCGGCAATGCTTAAGCGCTCGCTATGCTG 4274

Db 901 ttatctacacgacggggagtcaagaaactatgataagaaagaaatgaaagaaatgaa 960

Cp 4273 TTATCTACACGACGCGGAGTCAAGGTAATATGATGCAAGAAATAGATAGATGCTGACA 4214

Db 961 tagtgctcactgataaagcattgtaactgtaacactgtaacaaattttaaactatata 1020

Cp 4213 TAGTGCTCTCACTGATTAGCACTTGTAACTGCTGACCAACGTTTACTCATATATAATTT 4154

Db 1021 agattgatttaaaacttcatttttaatttaaaagatctagtgaaagactcttttttga 1080

Cp 4153 AGATTGATTAAACCTTCATTTTTTAATTTAAAGGATCTAGGTGAGGATCTTTTTTGATA 4094

Db 1081 atctcatgacccaaaatcccttaacgtgagtttttcttccactgaacactcaacactaa 1140

Cp 4093 ATCTCATGACCAAAATCCCTTAACGTGAGCTTTTTCGTTCCACTGAGCGTCAGACAGCTG 4044

Db 1141 aaaagatcaagaagatctcttgaagactcttttttctgagcgaatctgactgactgaaa 1200

Cp 4033 AAAAGATCAAGAGATCTCTTGAGATCTCTTTTTCGCTTAATCTGCTGCTGCTGCTG 9974

Db 1201 caaaaaaacaccgctaccagcggtgttttcttccagatcaagaaactcaacactcttt 1260

Cp 3973 CAATAAAACACCGCTTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9914

Db 1261 ttccgaaggtaactggcttcagcagcgcagataccaactactgtcttcttgaagtaag 1320

Cp 3913 TTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATAGCTGCTGCTTCTAGCTG 4854

Db 1321 cgttagtgacacacttcaagaactcttgagcagcgcacatacactcgtctgactaa 1380

Cp 3853 CGTAGTTAGGCCACCACTTCAAGAACTCTGTACACCGCTTACATACCTGCTGCTGCTG 4744

Db 1381 tctgttaccagtggtgctgccagtgccgataagctcgtctcttaccggttgaagtaag 1440

Cp 3793 TCGTGTACCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4744

Db 1441 gacgatagttaccgggataagcgacgctcggctcggctcggctcggctcggctcggctc 1500

Cp 3733 GACGATAGTTACCGGATAGGCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4674

Db 1501 ccagcttgagcgaacgacctacacgaactgaatcaccacagcgtgaagctatgaaaga 1560

Cp 3673 CCAGCTTGGAGCGAAGCGAGCTATACCGCAATCAGATACCTACCTGCTGCTGCTGCTG 4614

Db 1561 qcgccacgcttcccggaagggaagagcgaggtatcgggtaaagcgacgaagctcgaaga 1620

Cp 3613 GCGCCACGCTTCCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4564

Db 1621 cagagagcgacgagggagcttccaggggaaacgctcgtatcttattatgctcgtcga 1680

Cp 3553 CAGGAGAGCGACGAGGGAGGCTTCCAGGGGAAAGGCTGGTATGCTTATATATGCTGCTG 4494

Db 1681 ggttgcgcacctcacttgagcgtcgaatttttgaagctcgaagcgaagcgaagcga 1740

Cp 3493 GGTTCGCCACCTCTGACTTGGAGCGCTGAGATTTTGTGATGCTGCTGCTGCTGCTGCTG 4444

Db 1741 tatgaaaaacgcacgaacgagcgtttttacggttctgagcgttttgaagcgttttga 1800

Cp 3433 TATGAAAAACGCCACGCAACGGCGCTTTTACGGCTTCCTGCGCTTTTCTGCTGCTGCTG 4394

Db 1801 ctcaatgcttcttctcgaactataccctgattctgtaagaaacgtatctgagcgttttga 1860

Db 3228 gctttctcaatgctcagcgttagtagtctcagttcgggtgtaggtcgcttcctccaagct 3287
 QY 3612 GCTTTCTCATAGCTACGCTGCTAGGTATCTCAGTTTCGGGTAGGTCTCGTTCGCCICCAAGCT 3671
 Db 3288 gggctctgtgcacgaaccccccggttcagccgcagccgctgcgcttattccggttaactatcg 3347
 QY 3672 GGGCTGTGTGCAGCAACCCCGGCTTCAGCCGACCGGCTTCGGGCTTATCCGTAACACTCG 3731
 Db 3348 tcttgagtcacaccccggttaagcacagactatcgccactggcagcagccactgtaacag 3407
 QY 3732 TCTTGAGTCCACCCCGGTAAAGACAGACTATTCGCCACTGGCAGCAGCCACTGGTACAG 3791
 Db 3408 gattagcagagagaggtatgtatggcggtgtctacagagttcttgaagtggtggcctaacta 3467
 QY 3792 GATTAGCAGAGCGAGGTATGTAGGGGGTGCTACAGAGTCTTTGAAGTGGTGGCCCTAACTA 3851
 Db 3458 cgcctacactaaagacagatatttggtatctgcctctgtcgtcgtgaagcagtaacctcgg 3527
 QY 3852 CGGCTACACTAGAGACAGATATTTGGTATCTGCCGCTGTGCTGAAGCCAGTTACCTTCGG 3911
 Db 3528 aaaaagagttggttagctcttgatccggcgaacaaacccacccgctggtgagcgtggttttt 3587
 QY 3912 AAAAAAGTTGTTAGCTCTTATCCGGCAAAACAACACCCGCTGTTACGGGTGTTT 3971
 Db 3588 tgtttcagagcagcagatcacgcgcagagaaaaaagagatctcagaagatcctttgatctt 3647
 QY 3972 TGTTTCAAGCAGCAGATTACCGCGCAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 4031
 Db 3648 ttctcagggcttgacgtcagtgaaacaaacactcaagtttaagggtatttggtcatgag 3707
 QY 4032 TTTCTAGGGGCTTGACGCTCAGTGGAAACCAAACTACAGTTAAGGGATTTTGGTCATGAG 4091
 Db 3708 attatcaaaaagatcttcaccttagatccttttaattaaaaatgaatttttaaatcaat 3767
 QY 4092 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAAT 4151
 Db 3768 cttaatatatagtaaaactggtctgcagagttaccagttccttaactcagtgaggcacc 3827
 QY 4152 CTAAGTATATATAGTAACCTTGGTCTGACAGTTACCAATGCTTTAATCAGTGAGGCACC 4211
 Db 3828 tatctcagcagctctgtctatttcggttcacatagttcctgactgcccgtcggttagat 3887
 QY 4212 TATCTCAGCAGATCTGTCTATTTCGTTTCACTCCATAGTTGCTGACTCCCGGCTGTGTAGAT 4271
 Db 3888 aactacgatacggaggggcttaccatctgcccagctgctgcaatgatacgcgcagagacc 3947
 QY 4272 AACTAGCATACGGAGGGGCTTACCATCTGCCGCCAGTGTGCAATGATACATCCGAGACCC 4331
 Db 3948 acgctcacggctccagatttatcagcaataaaccagcagccggaagggcgagcgcag 4007
 QY 4332 ACCTCTACCGGGTCTCCAAATTTATCACTAAATAACAGCCAGCCGGAAGGCGGAG 4391
 Db 4008 aagtggtctcgcgaactttatccgcctccatccagttcttaattgttgcgggaagctag 4067
 QY 4392 AAGTGTCTCTGCAACTTTATCCGCTTCCATCCAGTCTATTATTGTTGCCGGGAAGCTAG 4451
 Db 4068 agtaactagttccagatgtaattgttgcgcacagctgttgcattgtcagagcagctgt 4127
 QY 4452 AGTAAGTAGTTCGCCAGTTAATAAGTTTGGCAACGCTTGTGCCATTGCTACAGGCATCGT 4511
 Db 4128 ggtgtcacgctcgtctgttggtatggtctcattcagctcgggttcccaacgataagcgc 4187
 QY 4512 GGTGTCACTGCTGCTGTGTGTGTATGCTTCAATTCAGTCCGGTTCCCAACGATCAAGGCG 4571
 Db 4188 agttacatgataccccctggttgaataaaagcgggttagctccttcggtcctccagatcgt 4247
 QY 4572 AGTTACATGATCCCGCATGTTGTGCAAAAAAAGCGGTAGTCTCTTCGTTCTCCCATGCT 4631
 Db 4248 tgcagaagtaagttggccgcagctggttataactcactgcttattgagcagcactgcataatc 4307
 QY 4632 TGTCAAGTAAGTTGGCCGACGTTTATCACTCATCTGTTATGGCAGCAGCTGCAATATC 4691
 Db 4308 tcttactgtcatgacatccgtaagatgctttctgtgactgtgagttactcaaccagatc 4367

QY 4692 TCTTACTGTGTCATGCGATCGTAAAGATGCTTTCTGTGACITGGTACITACCAACAGAC 4751
 Db 4368 attctgagaataagtagtgcgcggcagcagatctcttgcggcgctcaatacaggaataa 4427
 QY 4752 ATTCTGAGAATAAGTGTATGCGGCGACCGAGTTGTCTTGGCCGCGCTCAATACGGGATAA 4811
 Db 4428 taccgcgcacatagcagaactttaaaagtgctcatcattggaacagcttcttcgggagcg 4487
 QY 4812 TACCGCGTCACATACGACAACTTTAAAGTGTCTCATCTTGGAAAAAGTGTCTTCGGGGCG 4871
 Db 4488 aaaaactcagaagttcttaccgctgttgagatccaggttcgatgttaacccactcgtgcacc 4547
 QY 4872 AAAACTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 4931
 Db 4548 caactgatctcagcatcttttactttccaccagcgtttctgggtgagcaaaaaacacagaag 4607
 QY 4932 CAACGTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAACAGGAAG 4991
 Db 4608 gcaaatgcgcgcaaaaagggaataaagcgacacacgaaatgttgataactcatactctt 4667
 QY 4992 GCAAAATGCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGSAATACTCATACTCTT 5051
 Db 4668 ctttttcaatattattgaacatttatcaggggttatgtctctcagcagcagatcatt 4727
 QY 5052 CCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 5111
 Db 4728 taaatgtattgaaaaataaacaataaggaaggttcccgacacatttcccgaaaaatgcc 4787
 QY 5112 TGAATGATTATTGAAAAATAAACAATAAGGGTTTCCGGGACATTTCCCGAAAAAGTGGC 5171
 Db 4788 ac 4789
 QY 5172 AC 5173

RESULT 13
 ID Q84694 standard: DNA; 5042 BP.
 AC Q84694;
 DT 29-DEC-1995 (first entry)
 DE Plasmid glucoamylase Eco RI pGE.
 KW Neurospora crassa: glucoamylase gene; gla-1; promoter;
 KW expression construct; ss.
 OS Synthetic.
 PN WO9505474-A.
 PD 23-FEB-1995.
 PF 15-AUG-1994: G01789.
 PR 13-AUG-1993: GB-016883.
 PA (UYLE-) UNIV LEEDS.
 PI Parish JH, Radford A;
 DR WPI; 95-098771/13.
 PT Regulated glucoamylase promoter - useful for producing heterologous
 PT polypeptide(s) in filamentous fungi
 PS Claim 11; Fig 6; 36pp; English.
 CC The 5' primer (Q84690) encompassed the unique PvuII site at posn.
 CC 2163 of the N. crassa glucoamylase ORF and the 3' primer (Q84691)
 CC conformed an MroI site hybridises at the 3' end of the gla gene. The
 CC 5' upstream PCR fragment was amplified and cloned into the Sma I
 CC site in a pNEB 193 vector. The clone was named pMO. The remainder
 CC of the gla gene was inserted by digestion of the glucoamylase clone
 CC pGla-Xho I (Q84692). This plasmid contains the entire gla gene
 CC however the downstream unsequenced and non-transcribed area was
 CC deleted. pGla-Xho I was digested with Sac I and PvuII and the
 CC fragment ligated into the Sac I/PvuII sites of pMO I. The Sac I
 CC site of pGla-Xho I was derived from the linker and not from the
 CC coding region of glucoamylase consequently no glucoamylase
 CC sequence was deleted (see Q84693, pGla-Mro I). In an
 CC attempt to increase transcription efficiency, 1575 bp were deleted
 CC from the glucoamylase ORF, creating the plasmid pGE (plasmid
 CC glucoamylase, Eco RI). Deposits of plasmids pGla-Xho
 CC (Q84692), pGla-Mro I (Q84693), pGE (Q84694) have been made and
 CC the deposition details are to be added to the patent application.
 CC Sequence 5042 BP: 1291 A. 1307 C. 1229 G. 1225 T.


```

|||||
QY 5052 CTTTTTCAATATTATTGAAGCAITTTATCAGGGTTATTGCTCTAGCGGATACATATT 5111
|||||
Db 4978 tgaatatttgaagaaataaacaataagggttcgcgcacatttccccaaagtgc 5037
|||||
QY 5112 TGAATGATTAGAAAATAAACAATAAGGGTTCCGGGCAATTTCCCGAAAAGTGCC 5171
|||||
Db 5038 ac 5039
||
QY 5172 AC 5173

RESULT 14
ID Q40281 standard: DNA: 6926 BP.
AC Q40281:
DE Q2-AUG-1993 (first entry)
DE Sequence of clone pP2-gp160MN.
KW Plasmid; cloning; restriction site; HIV-1; MN-isolate; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..3529
FT /tag= a
FT /label= pS2gpt-P2 sequences
FT 2396..2851
FT /tag= b
FT /label= E. coli gpt gene
FT complement (3081..3323)
FT /tag= c
FT /label= vaccinia P7.5 promoter
FT 2396..2851
FT /tag= d
FT /label= P2 according to EP application Avipox
FT "intergenic region"
FT /note= "oligo P-artP(9)"
FT 6173..6926
FT /tag= e
FT /label= pN2gpt-S4
FT 3534..6001
FT /tag= f
FT /label= HIV-1 strain MN gp160 sequence (EMBL to
REH1VMNC)
FT AU9231269-A.
PN 04-MAR-1993.
PD 25-AUG-1992: 021269.
PF 25-AUG-1991: US-750080.
PR 20-JUL-1992: US-914738.
PA (TMMO ) IMMUNO AG.
P: Dorner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
OR WPL: 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example: Pages 177-180; 206pp; English.
CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. A new SmaI site
CC was introduced into the plasmid pN2gpt-S4 resulting in the plasmid
CC pS2gpt-S4. Subsequently the S4-promoter was exchanged by the P2-
CC promoter resulting in the plasmid pS2gpt-P2. For the modification
CC of the gp160-gene, a PCR-generated proximal fragment was exchanged
CC leading to a gp160-gene cassette with a minimal 5'-UTR.
SQ Sequence 6926 BP: 2064 A; 1470 C; 1686 G; 1706 T;
Query Match 42.5%; Score 2198; DB 7; Length 6926;
Best Local Similarity 99.9%; Pred No 0 one-00;
Matches 2100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 gtgcacatttttcggggaataatgctgcgcgaacccctatttttttataataacatt 60
|||||
Cp 5173 GTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTCATAATACATT 5114

```

```

Db 61 caaatatgtatccgctcatgagacaataaacctataaactcttcaataattgaaaa 120
|||||
Cp 5113 CAAATATGATCCGCTCATGAGACAATAACCTCATAAATGCTTCAATAATATGAAAA 5054
|||||
Db 121 ggaagatgatgattcaacatttccggtgctgccttattcccttttttgcgcatttt 180
|||||
Cp 5053 GGAAGAGATGAGTATTCAACATTTCCGTGTGCGCTTATTCCCTTTTTCGGGCAATT 4994
|||||
Db 181 gcttccctgttttgcctcaccccaaaaacgtggtgaaagtataaagatgctcaaatcagt 240
|||||
Cp 4993 GCCTTCCTGTTTTTGTCCACCCAGAAAGCGTGGTGAAGTAAAGATGCTGAAGATCAGT 4934
|||||
Db 241 tgggtgcacgagtggtttacatcgaactggaatcacaacagcggttaaatccttgaagt 300
|||||
Cp 4933 TGGGTGCACGAGTGCGTTACATCGAAGTGGATCTCAACACGCGTAAGATGCTTGAGATT 4874
|||||
Db 301 ttgcgccgaagaacgttttccaatgatgagcacttttaaaagtctctgtatgagcgcg 360
|||||
Cp 4873 TTGCGCCGGAAGACGTTTTTCCAATGATGAGCAGCTTTTAAAGTTCTGCTATGTGGCGCG 4814
|||||
Db 361 tattatccgtattgaccccggaagaagcaactcgatccgcgcatacactattctcaga 420
|||||
Cp 4813 TATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACATATTCICAGA 4754
|||||
Db 421 atgacttggttgagtaactaccagtcacagaaaagcatcttacgtagtggcatcacgtaa 480
|||||
Cp 4753 ATGACTTGGTTGAGTACTCACCGAGTCACAGAAAAGCATCTTACGGATGGCATGACAGATA 4694
|||||
Db 481 ggaattatgagtgctgctccataaccatgagtgataacactgcggcgaacttactctga 540
|||||
Cp 4693 GAGAATTATGCACTGCTGCCATAACCATGAGTGAATAACACTTGGCGGCAACTTACTTCTGA 4634
|||||
Db 541 caacgatcgagagaccgagagctaacccgcttttttgcacaacatggggatcatataa 600
|||||
Cp 4633 CAACGATCGGAGGACCGAAGAGAGCTAACCCGCTTTTTCACACACATGSGGATCAICATA 4574
|||||
Db 601 ctgccttgatcggttgggaacggagctgaatgaagccataaccacaacgacgagcgtaaca 660
|||||
Cp 4573 CTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCATACCAAAAGACGAGCGTGACA 4514
|||||
Db 661 ccacgatgcctgtagcaatggcaacaacgcttgcgcaaacattataactggcggaactactta 720
|||||
Cp 4513 CCACGATGCGCTGTAGCAATGGCAACACAGCTTGCGCAAACTATTAACTTGGGGAAGTACTTA 4454
|||||
Db 721 ctctagcttcccgccgaacaataatagactggatggagcggtataaagtgtcagagaccac 780
|||||
Cp 4453 CTCTAGCTTCCCGGCAACAATTATAGACTGGATGGAGCGGATAAAGTTGCAGAGACAC 4394
|||||
Db 781 ttctgcgtcgcccttcgcgctcggtgtgtttattgtgtataaatctggagccggtgaac 840
|||||
Cp 4393 TTCTGCGCTCGGCCCTTCGCGCTGCGTGTGTTTTATTGTGATAAAATCTGGAGCGGTGAGC 4334
|||||
Db 841 gtgggtctcgcggtatcattgacacactgggcccagatggtaaagccctccgctatcgtag 900
|||||
Cp 4333 GTGGGTCTCGCGGTATCATTTGACGACCTTGGGGCCAGATGGTAAAGCCCTCCCGGTATC 4274
|||||
Db 901 ttatctcacacgacggggagtcaggcaactatggatgaacgaatagacagactcgtcaga 960
|||||
Cp 4273 TTATCTACACGACGGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTCAGA 4214
|||||
Db 961 taggtgcctcactgattgaagcattggttaactgctcagacccaagttaactatataacttt 1020
|||||
Cp 4213 TAGGTGCGCTCAGCTGATTGAAGCATTTGTAACGTGTCAGACCAAGTTTACTCATATATATT 4154
|||||
Db 1021 agattgatttaaaacttcattttttaataaaagatctaggtaaaatcctttttgata 1080
|||||
Cp 4153 AGATTGATTAAAACTTTCATTTTAAAGAGATCTAGCTGAAGATCCCTTTTATATA 4034
|||||
Db 1081 atctcatgacaaaatcccttaacgtgagtttctgctccactgagcgtcagacccccctag 1140
|||||
Cp 4093 ATCTCATGACCAAAATCCCTTAAAGTGAGTTTTCGTTCCACTGAGTGTCAAGCCCCGAG 4034
|||||
Db 1141 aaagatcaaaagatctctcagagatccttttttctgcgcgttaactgtcgttgcaga 1200

```


Db 5006 gtggcacttttcggggaatgtgcggaacccctatttggtttttttataacatt 5065
 Cp 5173 GTGGCACTTTTCGGGAAATGTGCGGAACCCCTATTGTTTATTTTTCATAACATT 5114
 Db 5066 caaatatgtatccgctcatgagacaataaacctcgtataaatgtctcaataatagaaaa 5125
 Cp 5113 CAAATATGATCGGCTCATGAGACAATAAACCTCGTATAAATGCTCAATAATATTGAANA 5054
 Db 5126 ggaagagtatgattatcaacatttccgtgtgcgccttatcccttttttgcggcatttt 5185
 Cp 5053 GGAAGAGTATGAGTATTCAACATTTCGGTGTGCGCCATTATCCCTTTTTCGGGCATTIT 4994
 Db 5186 gccttctgtttttgtcaccaccgaaacgctggtgaaagtataagatcgtgaagatcagt 5245
 Cp 4993 GCTTTCCTGTTTGTCTACCCAGAAACGCTGGTGAAGATGAAGATGCTGAAGATCAGT 4934
 Db 5246 tgggtcacagatgggttacatcgaactgagatctcaacagcggttaagatccttgagtt 5305
 Cp 4933 TGGTGGACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTGAAGATCCTTGAGAGTT 4874
 Db 5306 ttgcgcgcgaagaacattttccaatgataagcaactttttaaagtctgtatgtggcgcg 5365
 Cp 4873 TTGCCCCGAAGAAGCTTTTCCAATGATGAGCACITTTTAAAGTTCTGTATGTGGCGCG 4814
 Db 5366 tatatccctatgtgacgcgcggaagcaactcgtgcgcgataacattctcaga 5425
 Cp 4813 TATATCCCGTATTTGACGCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGA 4754
 Db 5426 atgacttggttgactcaccagtcacagaaagacatcttacgtaggcatgacatgaa 5485
 Cp 4753 ATGACTTGGTTGAGTACTTACCAGTCTACAGAAAGCATCTTACGGATGGCATGACAGTAA 4694
 Db 5486 gagaattatcagctgctgcataaacatagtgataaacactgcgcgaacttaactcttga 5545
 Cp 4593 GAGAATTATCGACTGTGCTAATCAATGATGATGATGATGATGATGATGATGATGATGAT 4634
 Db 5546 caacgactgagagacgaagagactaacctgtttttgcacaaacatgggggatcatgtaa 5605
 Cp 4533 CAACGATGAGAGACGCAAGAGCTAAGGAGCTAAGGAGCTAAGGAGCTAAGGAGCTAAG 4574
 Db 5606 ctgccttgatcgttgggaacgagctgaatgaagcacaataccaaacgacgagcgtgaca 5665
 Cp 4573 CTGCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACACGAGCGTGACA 4514
 Db 5666 ccaacgatgctgtagcaatggaacacacttgcgcaaacatttaactggcaactactta 5725
 Cp 4513 CCACGATGCGCTGAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTACTTA 4454
 Db 5726 ctctagcttcccggaacaataatagactgtagatgagagcgatataaagtgcaggaccac 5785
 Cp 4453 CTCATATTCCTGCGGTATCATATGCAACATGAGATGAGATGAGATGAGATGAGATGAG 4394
 Db 5786 ttctgcctgcgccttcccggtgctggtgtttattgctgataaatctggagcgggtgagc 5845
 Cp 4393 TTCTGCGCTGCGGCTTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4334
 Db 5846 gtgggtctcggatcatcattgagcactggggcgagatggtgaagccctcccgatcgtag 5905
 Cp 4333 GTGGGTTCCTGCGGTATCATATGCAACATGAGATGAGATGAGATGAGATGAGATGAG 4274
 Db 5906 ttatctacacgagggagtcaggaactatgagtaaacgaataagacagatcgtcgtaga 5965
 Cp 4273 TTATCTACACGAGGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGA 4214
 Db 5966 tagtgctcactgattaaagcattggttaactgtcacacaaagttaactcatataacttt 6025
 Cp 4213 TAGTGCTCTCAGTATTAAAGCATTTGTAAGTGTACAGCAAGCTTTACATATATATCTTT 4154
 Db 6026 agattgatttaaaccttcatttttaatttaaaagatcagtggaagatcctttttgata 6085
 Cp 4153 ACATTGATTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATA 4094
 Db 6086 atctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtag 6145

Cp 4093 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCTCAGACCCGCTAG 4034
 Db 6146 aaaaatcaaaagatctctcttgagatcccttttttctgcgcgtaaatctgctgtctgcaaa 6205
 Cp 4033 AAAAGATCAAAAGATCTCTTTTATGATGCTTTTCTTCTGCGGTAAATTTGCTGTGAAA 3974
 Db 6206 caaaaaaaccccgctacacgctgtgtgtttgttgcgcggtacaaagactaccactttt 6265
 Cp 3973 CAAAAAACCCACCGTATACGAGGCTGCTTTGTTTGTGATCAAGATCAAGATCAAGT 3914
 Db 6265 ttccgaagatcaactggtcttcgcgcgagcgagatcaacaaatctctctctctctagc 6325
 Cp 3913 TTCCGAAGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGCTCTCTAGTGAGC 3854
 Db 6326 cgtatgtaggccacacttcaagaactctgtagcaccgcctacatacctcctctgtaa 6385
 Cp 3853 CGTAGTTAGGCTCACCACCTTCAAGAACTCTGTAGCAGCGCTTACATAGCTCTGCTTGT 3794
 Db 6386 tectgttaccagtgctgctccagtgcgataaagtgtgttaccgggttgagactcaa 6445
 Cp 3793 TCTGTATTACAGTGCTGCTGCTCAGTGCGGATAGTGTGTGCTTTACCGGCTTGGACTCA 3734
 Db 6446 gacgataattaccgataaagcgagcgctcggctgaacgggggttcgtacacacagc 6505
 Cp 3733 GACGATAGTTACGGATTAAGCGGACAGCGCTGCGCTGAACGGGGGTTCTGTACATAGC 3674
 Db 6506 ccagcttgagcggaacacctacacccgaactgagatacctacacgtagcgtatcaaaa 6565
 Cp 3673 CCAGCTTGGAGCGAAGACCTTACACCGAACTTGAATACCTACACGCTGAGCTATGAGAAA 3614
 Db 6566 ggcacacgcttcccgaaagggagaaagcgagcgatcccgtaagcgagcgaggtcgaa 6625
 Cp 3613 GCGCCACGCTTCCCGAAGGGAGAAAGGCGACAGTATCCGCTAAGCGGAGGGTCCGAA 3554
 Db 6625 caggagcgacgagcgagagcttccaggggaaacgcctgatatctttatagctctcg 6685
 Cp 3553 CAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGTCTCTGCG 3494
 Db 6686 ggttcgcacactctgacttgagctgatatcttctgtagctcgtcagggggcgagcgc 6745
 Cp 3493 GGTTCGCTACCTCTGACTTGAAGCTGCTGATTTTGTGATGCTGCTGAGGCGGAGG 3434
 Db 6746 tatgaaaaacgccagcaacgcgcttttaccggttccgcttttctgctgctttt 6805
 Cp 3433 TATGAAAAACGCGACAGCAAGCGGCTTTTAAAGGTTCTTCTGATTAAGCGCTTIG 3374
 Db 6806 ctacatgttttctcgtgattatccctgattctctgagataaacgattaccgctttg 6865
 Cp 3373 CTCACATGTTCTTCTGCTATCCCTGATTTCTGATTAAGCGCTTATTAAGCGCTTIG 3314
 Db 6866 agtgagctgataccgctcgcgcagccgaacgagcgagcgagctcaatgagcgagc 6925
 Cp 3313 AGTGAGCTGATACCGCTCGCGCAGCCGAACGAGCGAGCGAGGAGTCACTGAGCGAGG 3254
 Db 6926 aagcgaagagcgcccaataagcaaacccgctctcccgcgcggttgagcgatcattaat 6985
 Cp 3253 AAGCGGAAGAGCGCCCAATACGCAAAAGCGCTCTCCCGCGCGGTGGCGGATTCATTAAT 3194
 Db 6986 gcaagtcgcaacagaggtttcccgactggaagcgagcgagcgagcgagcgagcgagcgag 7045
 Cp 3193 GAGCTGCGACGACAGGTTTCCCGACTGTGAAAGCGGCGAGTGAAGCGCAATTAATG 3134
 Db 7046 tgaattagctcactcattaggaaccccgaggttttacactttatgcttcgagctcgtatgt 7105
 Cp 3133 TGAGTTAGCTCATTATAGGCAACCGAGCTTTTACACTTTTACTTCTTCCGCTCTATGT 3074
 Db 7106 tctgtggaattgtgagcgataacaatttcacagaggaaacagctatgacatgacag 7165
 Cp 3073 TGTGTGAAATTTGAGCGGATAAACAATTTCAACAAGAAAGATATGACCATGATTACG 3014
 Db 7166 ccaagctcgaattaacctcactaaagggaacaaagcctcg 7207

CP 3013 CCAAGGGGCAATTAAAGCTTCTACTAAGGGAACAAAGCTGG 2972

Search completed: Tue Mar 16 12:20:37 1999
Job time : 915 secs.

 WAREHOUSE
 ***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

Mpsrch_nu n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 16 09:49:15 1999; Maspar time 6574.27 Seconds
 Tabular output not generated. 1409.082 Million cell updates/sec

Title: >US-09-020-716-7

Description: (1-5173) from US09020716.seq

Perfect Score: 5173

N.A. Sequence:

Comp: 1 CTAATTGTAAGCGTTAATA... ATTTCGGGAAAGTGCAC 5173

GATTTAACATTGCAATTAT... TAAAGGGGCTTTTCACGGTG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: embl-est56

Database: genbank-est109

1:em-est1 2:em-gss1 3:em-gss2 4:em-gss3

5:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13

10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17

14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21

18:gb-est21 19:gb-est3 20:gb-est4 21:gb-est5 22:gb-est6

23:gb-est7 24:gb-est8 25:gb-est9 26:gb-gss1 27:gb-gss2

28:gb-gss3 29:gb-gss4

Statistics: Mean 12.749; Variance 2.640; scale 4.830

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	564	12.8	711	26	AG004508	Homo sapiens genomic D	0.00e+00
2	564	12.8	711	27	AG010947	Homo sapiens genomic D	0.00e+00
3	557	12.7	721	27	AG011001	Homo sapiens genomic D	0.00e+00
4	557	12.7	721	26	AG004562	Homo sapiens genomic D	0.00e+00
5	535	12.3	718	26	AG004363	Homo sapiens genomic D	0.00e+00
6	535	12.3	718	27	AG010489	Homo sapiens genomic D	0.00e+00
7	530	12.2	698	26	AG003787	Homo sapiens genomic D	0.00e+00
8	530	12.2	598	27	AG009976	Homo sapiens genomic D	0.00e+00
9	524	12.1	596	27	AG009765	Homo sapiens genomic D	0.00e+00
10	524	12.1	596	26	AG003376	Homo sapiens genomic D	0.00e+00
11	524	12.1	720	28	AG013858	Homo sapiens genomic D	0.00e+00
12	524	12.1	720	26	AG006062	Homo sapiens genomic D	0.00e+00
13	619	12.0	747	26	AG007052	Homo sapiens genomic D	0.00e+00

C	14	609	11.8	644	28	AQ115013	CIT-HSP-2374C24.TF CIT	0.00e+00
C	15	603	11.7	692	27	AG0040352	CIT-HSP-2327K21.TF CIT	0.00e+00
C	16	603	11.7	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	598	11.6	748	17	A1124281	1.53 Drosophila 8-12 h	0.00e+00
C	18	593	11.5	640	27	AG074298	CIT-HSP-2382L21.TF CIT	0.00e+00
C	19	595	11.5	654	26	B78704	CIT-HSP-2382L21.TF CIT	0.00e+00
C	20	594	11.5	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	21	597	11.5	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	22	588	11.4	652	28	AQ108633	CIT-HSP-2379M3.TF CIT	0.00e+00
C	23	583	11.3	594	26	C17B6	CIT-HSP-2379M3.TF CIT	0.00e+00
C	24	583	11.3	566	27	AG079096	CIT-HSP-2356E24.TF CIT	0.00e+00
C	25	577	11.2	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.9	642	28	AQ113213	CIT-HSP-2372L22.TF CIT	0.00e+00
C	27	567	10.8	682	27	AG074693	CIT-HSP-2301L23.TF CIT	0.00e+00
C	28	556	10.8	615	28	AQ113424	CIT-HSP-2376F2.TF CIT	0.00e+00
C	29	559	10.8	742	28	AQ108392	CIT-HSP-2383E2.TF CIT	0.00e+00
C	30	551	10.7	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.6	634	27	AG0038010	CIT-HSP-2326F2.TF CIT	0.00e+00
C	32	549	10.5	653	27	AG0077504	CIT-HSP-2365K21.TF CIT	0.00e+00
C	33	544	10.5	633	27	AG040787	CIT-HSP-2365K22.TF CIT	0.00e+00
C	34	545	10.5	661	27	AQ057603	CIT-HSP-2345M2.TF CIT	0.00e+00
C	35	545	10.5	690	27	AQ0029399	RPC111-39F1.TP RPC111	0.00e+00
C	36	536	10.4	626	28	AQ111342	CIT-HSP-2371M2.TF CIT	0.00e+00
C	37	540	10.4	637	27	AQ0077360	CIT-HSP-2365E1.TF CIT	0.00e+00
C	38	531	10.3	549	27	AQ003191	RPC111-19E5.TP RPC111	0.00e+00
C	39	534	10.3	594	26	B71605	RPC111-9N1.TP RPC111 H	0.00e+00
C	40	526	10.2	540	27	AQ074307	CIT-HSP-2382M1.TF CIT	0.00e+00
C	41	526	10.2	594	26	AS007826	Homo sapiens genomic D	0.00e+00
C	42	527	10.2	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C	43	530	10.2	630	27	AQ003238	RPC111-19L13.TP RPC111	0.00e+00
C	44	530	10.2	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C	45	525	10.1	500	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00

ALIGNMENTS

RESULT 1 AG004508 711 bp DNA GSS 29-JAN-1998
 LOCUS Homo sapiens genomic DNA, 21q region, clone: S594BG38, genomic
 DEFINITION survey sequence.
 ACCESSION AG004508
 NID G8222058
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: S594BG38.
 ORGANISM Homo sapiens
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia, Eutheria;
 Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 711)
 AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 711)
 AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JS-
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
 (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
 Fax: 0427-78-9561)

FEATURES
 Location/Qualifiers
 1..711
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="S594BG38"
 /map="21q"

BASE COUNT 169 a 188 c 172 g 182 t

Query Match 12.8%; Score 664; DB 26; Length 711;

Best Local Similarity 99.6%; Pred. No. 0.00e+00;

Matches 68%; Conservative 0; Mismatches 0; Indels 3; Caps 3;


```

REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1998) to the DBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES             Location/Qualifiers
     source           1..721
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="879G12X91"
                     /map="21q"
BASE COUNT  195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match      12.7%: Score 657; DB 27; Length 721;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAATTATGTCAGTGTGTCATACCATGAGTGTAT 91
Cp 4718 CATCTTACGGATGGCATGACAGTAAGAGAATTATGTCAGTGTGTCATACCATGAGTGTAT 4659
Db 92 AACACTCGCGCCAACTTACTTCTGACAGCATGCGGAGGACCGAAGAGCTAACCGCTTTT 151
Cp 4658 AACACTCGCGCCAACTTACTTCTGACAGCATGCGGAGGACCGAAGAGCTAACCGCTTTT 4599
Db 152 TTGCACAAATGGGGATGATGTAAGTACGCTGTGCTGTTGGAAACCGAGCTGAATGAA 211
Cp 4598 TTGCACAAATGGGGATGATGTAAGTACGCTGTGCTGTTGGAAACCGAGCTGAATGAA 4539
Db 212 GCTATACCAACACGAGCGGTGACACGATGCTGTAGCAATGGCAACACGTTGGCG 271
Cp 4538 GCTATACCAACACGAGCGGTGACACGATGCTGTAGCAATGGCAACACGTTGGCG 4479
Db 272 AAACATTTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG 331
Cp 4478 AAACATTTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG 4419
Db 332 GAGCGGATAAAGTTGCAGGACCACTTCTCGCTCGGCCCTTCCGGCTGGCTGGTTTATT 391
Cp 4418 GAGCGGATAAAGTTGCAGGACCACTTCTCGCTCGGCCCTTCCGGCTGGCTGGTTTATT 4359
Db 392 GCTGATAAATCTGGAGCCGGTGAAGTGGTCTCGCGGTATCATTTGCAGCACTGGGGCCA 451
Cp 4358 GCTGATAAATCTGGAGCCGGTGAAGTGGTCTCGCGGTATCATTTGCAGCACTGGGGCCA 4299
Db 452 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAGCGGAGTCAGGCAACTATGATG 511
Cp 4298 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAGCGGAGTCAGGCAACTATGATG 4239
Db 512 GAACGAATATAGACAGATCGGTGAGATAGTGGCTCACTGATTAAAGC-TTGGTAACGTGCA 570
Cp 4238 GAACGAATATAGACAGATCGGTGAGATAGTGGCTCACTGATTAAAGC-TTGGTAACGTGCA 4179
Db 571 GACCAAGTTTACTC-TATATACCTTTTATAGATGATTTAAACCTTCATTTTAAATGAGG 629
Cp 4178 GACCAAGTTTACTC-TATATACCTTTTATAGATGATTTAAACCTTCATTTTAAATGAGG 4119
Db 630 ATCTAGGTGAAGATCCTTTT-GATAATCNCATGACCAAA-TCCCTTAAGCTGAGTTTCG 687
Cp 4118 ATCTAGGTGAAGATCCTTTT-GATAATCNCATGACCAAA-TCCCTTAAGCTGAGTTTCG 6059
Db 688 TTCCACTGAGCGTCAGACCCCGTAGAAAAAT 719
Cp 4058 TTCCACTGAGCGTCAGACCCCGTAGAAAAAT 4027

```

```

4
RESULT AG004662 721 bp DNA GSS 30-JAN-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic
DEFINITION survey sequence.
ACCESSION AG004662
NID Q2826191
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:879G12X91.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-1998) to the DBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES             Location/Qualifiers
     source           1..721
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="879G12X91"
                     /map="21q"
BASE COUNT  195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match      12.7%: Score 657; DB 26; Length 721;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAATTATGTCAGTGTGTCATACCATGAGTGTAT 91
Cp 4718 CATCTTACGGATGGCATGACAGTAAGAGAATTATGTCAGTGTGTCATACCATGAGTGTAT 4559
Db 92 AACACTCGCGCCAACTTACTTCTGACAGCATGCGGAGGACCGAAGAGCTAACCGCTTTT 151
Cp 4658 AACACTCGCGCCAACTTACTTCTGACAGCATGCGGAGGACCGAAGAGCTAACCGCTTTT 4599
Db 152 TTGCACAAATGGGGATGATGTAAGTACGCTGTGCTGTTGGAAACCGAGCTGAATGAA 211
Cp 4598 TTGCACAAATGGGGATGATGTAAGTACGCTGTGCTGTTGGAAACCGAGCTGAATGAA 4539
Db 212 GCTATACCAACACGAGCGGTGACACGATGCTGTAGCAATGGCAACACGTTGGCG 271
Cp 4538 GCTATACCAACACGAGCGGTGACACGATGCTGTAGCAATGGCAACACGTTGGCG 4479
Db 272 AAACATTTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG 331
Cp 4478 AAACATTTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG 4419
Db 332 GAGCGGATAAAGTTGCAGGACCACTTCTCGCTCGGCCCTTCCGGCTGGCTGGTTTATT 391
Cp 4418 GAGCGGATAAAGTTGCAGGACCACTTCTCGCTCGGCCCTTCCGGCTGGCTGGTTTATT 4359
Db 392 GCTGATAAATCTGGAGCCGGTGAAGTGGTCTCGCGGTATCATTTGCAGCACTGGGGCCA 451
Cp 4358 GCTGATAAATCTGGAGCCGGTGAAGTGGTCTCGCGGTATCATTTGCAGCACTGGGGCCA 4299
Db 452 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAGCGGAGTCAGGCAACTATGATG 511
Cp 4298 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAGCGGAGTCAGGCAACTATGATG 4239
Db 512 GAACGAATATAGACAGATCGGTGAGATAGTGGCTCACTGATTAAAGC-TTGGTAACGTGCA 570

```



```

Db 141 CATCAAAATCGAGCTCAAGTTCAGAGTGGCGAAACCCGACAGGACTATAAGATAC 200
      |||||||
Qy 3454 CATCAAAATCGAGCTCAAGTTCAGAGTGGCGAAACCCGACAGGACTATAAGATAC 3513
      |||||||
Db 201 CAGCGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCTCTGTTCCGACCTCCGCTTACC 260
      |||||||
Qy 3514 CAGCGTTTCCCGCTGGAAGCTCCCTGCTGGCTCTCTCTGTTCCGACCTCCGCTTACC 3573
      |||||||
Db 261 GGATAGCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTCATAGCTCAGCTGT 320
      |||||||
Qy 3574 GGATAGCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTCATAGCTCAGCTGT 3633
      |||||||
Db 321 AGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGCAGGAACCCGCC 380
      |||||||
Qy 3634 AGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGCAGGAACCCGCC 3693
      |||||||
Db 381 GTTCAGCGGACGCTGGCGCTTATCCGGTAACTATCTGCTTGTAGTCCAAACCCGGTAA 440
      |||||||
Qy 3694 GTTCAGCGGACGCTGGCGCTTATCCGGTAACTATCTGCTTGTAGTCCAAACCCGGTAA 3753
      |||||||
Db 441 CAGCACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGATTAGCAGAGCNAGGTATGTA 500
      |||||||
Qy 3754 CAGCACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGATTAGCAGAGCNAGGTATGTA 3813
      |||||||
Db 501 GCGCGTGTCAACAGAGTCTTGAAGTGGTGGCTAACTACGGCTAACTAGAGAAGACAG 560
      |||||||
Qy 3814 GCGCGTGTCAACAGAGTCTTGAAGTGGTGGCTAACTACGGCTAACTAGAGAAGACAG 3871
      |||||||
Db 561 TATTTGATCTGCGCTCTGCTGAGCCAGATTANCTTTCGNAAGAGAGTGGTAGCTCT 620
      |||||||
Qy 3872 TATTTGATCTGCGCTCTGCTGAGCCAGATTANCTTTCGNAAGAGAGTGGTAGCTCT 3930
      |||||||
Db 621 TGATCGGGCAAAACCAACCCGCTGTAGCG-TGTTTTTTTGGTTCGAAGCAGCAGATA 679
      |||||||
Qy 3931 TGATCGGGCAAAACCAACCCGCTGTAGCGTGGTGTGTTTTTGGTTCGAAGCAGCAGATT 3990
      |||||||
Db 680 ACCGGCAGAAAAAAGATCTCAAGA-GATC-TTTGATCTT 718
      |||||||
Qy 3991 ACCGGCAGAAAAAAGATCTCAAGAAGATCTCTTTGATCTT 4031
      |||||||

RESULT 7
LOCUS AG003787 698 bp DNA GSS 07-JAN-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
sequence.
ACCESSION AG003787
NID 92754689
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T485XN.
ORGANISM Homo sapiens
REFERENCE Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Primates; Catarrhini; Hominidae; Homo.
TITLE Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
JOURNAL 1 (bases 1 to 698)
PUBLISHED Only in Database (1998) In press
REFERENCE 2 (bases 1 to 698)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
JOURNAL Direct Submission
TITLE Submitted (07-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp. Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
Location/Qualifiers
source
1..698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="T485XN"
/map="21q"
BASE COUNT 189 a 153 c 171 g 175 t 10 others
ORIGIN

```

```

Query Match 12.2% Score 630; DB 26; Length 698;
Best Local Similarity 99.0%; Pred.No.0.00e-00; Mismatches 0; Indels 5; Gaps 5;
Matches 661; Conservative 0;

Db 31 CTATCGGATGGCATGACAGTAAGAGAATTATGCAAGTCTGCCATAACCATGAGTATAAC 90
      |||||||
Cp 4715 CTATCGGATGGCATGACAGTAAGAGAATTATGCAAGTCTGCCATAACCATGAGTATAAC 4656
      |||||||
Db 91 ACTGGGGTCAACTTACTTCTGACAAACGATCGAGAGCCGGAAGAGCTTAACCGCTTTTGG 150
      |||||||
Cp 4655 ACTGGGGTCAACTTACTTCTGACAAACGATCGAGAGCCGGAAGAGCTTAACCGCTTTTGG 4596
      |||||||
Db 151 CACAACATGGGGGATCATGTAACTCGCTTGCATGTTGGGAACCGGAGCTGAATGAAGCC 210
      |||||||
Cp 4595 CACAACATGGGGGATCATGTAACTCGCTTGCATGTTGGGAACCGGAGCTGAATGAAGCC 4536
      |||||||
Db 211 ATACCAACGAGGAGCGTGCACACCATGCTGTAGCAATGGCAACAGCTTGGCGAAA 270
      |||||||
Cp 4535 ATACCAACGAGGAGCGTGCACACCATGCTGTAGCAATGGCAACAGCTTGGCGAAA 4476
      |||||||
Db 271 CTATTAACTGGGCACTACTTACTTACTTCCCGGCAACAATTAATAGACTGATGGAG 330
      |||||||
Cp 4475 CTATTAACTGGGCACTACTTACTTACTTCCCGGCAACAATTAATAGACTGATGGAG 4416
      |||||||
Db 331 GCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCTTCCGGCTGGCTGGTTTATGCT 390
      |||||||
Cp 4415 GCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCTTCCGGCTGGCTGGTTTATGCT 4356
      |||||||
Db 391 GATAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATCATTTGAGCACTGGGCCAGAT 450
      |||||||
Cp 4355 GATAAATCTGGAGCGGTGAGCGTGGGTCTCGCGGTATCATTTGAGCACTGGGCCAGAT 4296
      |||||||
Db 451 GGTAGCCCTCCGCTATCTAGTATCTACAGCAGCGG-AGTCAGGCAACTATGGATGA 509
      |||||||
Cp 4295 GGTAGCCCTCCGCTATCTAGTATCTACAGCAGCGGAGTCAGGCAACTATGGATGA 4236
      |||||||
Db 510 CGAATAGACAGATTTCGCTGAGTAGTTCCTCACTGATTAAAGCATTTGTAACCTGTGCA 569
      |||||||
Cp 4235 CGAATAGACAGATTTCGCTGAGTAGTTCCTCACTGATTAAAGCATTTGTAACCTGTGCA 4177
      |||||||
Db 570 CCAAGTTTACTC-TATATCTTACTTATGATTAAACCTTCATTTTAAATTAAGAT 628
      |||||||
Cp 4176 CCAAGTTTACTC-TATATCTTACTTATGATTAAACCTTCATTTTAAATTAAGAT 4117
      |||||||
Db 629 CTAGGTGAAGATCTTTT-GATAATCTCATGCAAAATTCCTTAACGTGAGTTTGGT 687
      |||||||
Cp 4116 CTAGGTGAAGATCTTTT-GATAATCTCATGCAAAATTCCTTAACGTGAGTTTGGT 4057
      |||||||
Db 688 TCCACTGA 695
      |||||||
Cp 4056 -CCACTGA 4050

RESULT 8
LOCUS AG009976 698 bp DNA GSS 05-JUL-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey
sequence.
ACCESSION AG009976
NID 93289962
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T485XN.
ORGANISM Homo sapiens
REFERENCE Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Primates; Catarrhini; Hominidae; Homo.
TITLE Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL 1 (bases 1 to 698)
PUBLISHED Only in Database (1998) In press
REFERENCE 2 (bases 1 to 698)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL Direct Submission
TITLE Submitted (06-JUL-1998) to the DDBJ/EMBL/GenBank databases.

```



```

/chrmosome="21"
/clone="c48B2A23"
/map="21q"

BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 12.0%; Score 619, DB 26, Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTGGGCTCTGCTGAAGCAGTATCTTCGAAAAAGAGTGTAGTCTTGTATCGGC 100
QY ATCTGGGCTCTGCTGAAGCAGTATCTTCGAAAAAGAGTGTAGTCTTGTATCGGC 3939
Db 101 AAACAACACCGCTCGTAGCGGTG-TTTTTTTTGTTCAGCAGCAGATACCGGAGA 159
QY AAACAACACCGCTCGTAGCGGTGTTTTTTTGTTCAGCAGCAGATACCGGAGA 3999
Db 160 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAAC 219
QY AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAAC 4059
Db 220 GAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAAAGGATCTTCACTAGATC 279
QY GAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAAAGGATCTTCACTAGATC 4119
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGATCAAACTTGGTCT 339
QY CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGATCAAACTTGGTCT 4179
Db 340 GACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGCTATTTCGTCA 399
QY GACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGCTATTTCGTCA 4239
Db 400 TCCATAGTTCCTGACTCCCGCTGCTAGATCAATCAGTACGAGGAGGT-ACCATCT 458
QY TCCATAGTTCCTGACTCCCGCTGCTAGATCAATCAGTACGAGGAGGTTCACATCT 4299
Db 459 GGGCCAGTCTGCAATGATAGCGGAGAGCCAGCTACCGGCTCCAGATTTATCAGCA 518
QY GGGCCAGTCTGCAATGATAGCGGAGAGCCAGCTACCGGCTCCAGATTTATCAGCA 4359
Db 519 ATAAACAGCAGCCGAGGAGGCGGAGGCA-AGTGTGCTCGCAACTTTATCGGCTC 577
QY ATAAACAGCAGCCGAGGAGGCGGAGGCGGAGGAGGCTCGCTGCAACTTTATCGGCTC 4418
Db 578 CATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAAGTAGTTTCCGCAAGTTAATAGT 637
QY CATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAAGTAGTTTCCGCAAGTTAATAGT 4476
Db 638 TTGGCAACGTTGTTGNCATGCTGAGGAGATGTTGTTGTCACGNCCTCCNTTGGGTAA 697
QY TTGGCAACGTTGTTGNCATGCTGAGGAGATGTTGTTGTCACGNCCTCCNTTGGGTAA 4535
Db 698 GGTTCATTTCAAGTCCGGTCCCAACGATCAAGGGGTTAACATGAT 746
QY GGTTCATTTCAAGTCCGGTCCCAACGATCAAGGGGTTAACATGAT 4582

RESULT 14
LOCUS A0115013 644 bp DNA GSS 28-AUG-1998
DEFINITION CIT-HSP-2374C24.FR CIT-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION A0115013
NID g3491134
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wibic, C., Shizuya, H., Simon, M. and

```

```

Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
Location/Qualifiers
1..644
/organism="Homo sapiens"
/clone="2374C24"
/sex="Male"
/cell_type="Sperm"
135 a 183 c 145 g 179 t 2 others

BASE COUNT 135 a 183 c 145 g 179 t 2 others
ORIGIN

Query Match 11.8%; Score 609, DB 28, Length 644;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCACCTGGCGCTCGTTTTACACAGCTGCTGACTGGGAAACCCCTGCGTTACCAACITTAAT 64
Cp 618 TCACCTGGCGCTCGTTTTACACAGCTGCTGACTGGGAAACCCCTGCGTTACCAACITTAAT 559
Db 65 CGCCTTGACGACATCCCGCTTTCCGCGAGCTGGGTAAATAGGAAGAGGCGCGACCCGAT 124
Cp 558 CGCCTTGACGACATCCCGCTTTCCGCGAGCTGGGTAAATAGGAAGAGGCGCGACCCGAT 499
Db 125 CGCCTTCCCAACAGTTGCGAGCCTGAATGGGAATGG-ACGGCGCCTGTAGCGGGGCA 183
Cp 498 CGCCTTCCCAACAGTTGCGAGCCTGAATGGGAATGGGACGCGCCTGTAGCGGGGCA 439
Db 184 TTAAGCGCGCGGGTGTGGTGTACGGCGAGCGTGACCGGTACACTTCCAGCGGCCA 243
Cp 438 TTAAGCGCGCGGGTGTGGTGTACGGCGAGCGGTGACCGGTACACTTCCAGCGGCCA 379
Db 244 GCGCCGCTCCCTTTCCGCTTTCTCCCTTCTCCCTTCTCGGCAAGTTGCGGTGCTTTCCGCT 303
Cp 378 GCGCCGCTCCCTTTCCGCTTTCTCCCTTCTCCCTTCTCGGCAAGTTGCGGTGCTTTCCGCT 319
Db 304 CAACTCTAAATCGGGGCTCCCTTTAGSGTTCGGATTTAGTGTTTTACGACCTGAC 363
Cp 318 CAACTCTAAATCGGGGCTCCCTTTAGSGTTCGGATTTAGTGTTTTACGACCTGAC 259
Db 364 GCGAAAAACTTGAATAGGTCATGTTACGTCAGTGTGGCCATCGCCCTGATACAGCTT 423
Cp 258 GCGAAAAACTTGAATAGGTCATGTTACGTCAGTGTGGCCATCGCCCTGATACAGCTT 199
Db 424 TTTGCGCCTTTGACGTTGGAGTCCACGTTCTTTTAAATAGTGACCTCTTGTTCACAACTGGA 483
Cp 198 TTTGCGCCTTTGACGTTGGAGTCCACGTTCTTTTAAATAGTGACCTCTTGTTCACAACTGGA 139
Db 484 ACAACTCAACCTTATCTCGGTCTATTCTTTGATTTATAGGATTTTCCGATTTCC 543
Cp 138 ACAACTCAACCTTATCTCGGTCTATTCTTTGATTTATAGGATTTTCCGATTTCC 79
Db 544 GCCTATTGTTAAAAAATGAGCTGATTTAAACAAAAATTTAACGGGAATTTTAAACAAATA 603
Cp 78 GCCTATTGTTAAAAAATGAGCTGATTTAAACAAAAATTTAACGGGAATTTTAAACAAATA 19

```

```

Db 504 ITACGCTTACAAATTT 619
CP 18 ITACGCTTACAAATTT 3

RESULT 15 A0040452 692 bp DKA GSS 10-JUL-1998
LOCUS CIT-HSP-2327K21.TF CIT-HSP Homo sapiens genomic clone 2327K21,
DEFINITION genomic survey sequence.
ACCESSION A0040352
NID g3306384
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria.
Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 692)
Adams M D, Pounsley S D, Field C E, Bass S, Linher K,
Golden K, Berry K, Granter D, Suh E, Wible C, Shizuya H,
Simon M, and Venter J C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Other_GSSs: CIT-HSP-2327K21.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES             Location/Qualifiers
     source            1..692
     /organism="Homo sapiens"
     /note="Vector: pReloBAC11; Site_1: HindIII; Site_2:
HindIII"
     /db_xref="taxon:9606"
     /clone="2327K21"
     /clone_lib="CIT-HSP"
     /sex="Male"
     /cell_type="Sperm"
BASE COUNT 158 a 193 c 186 g 155 t
ORIGIN

Query Match 11.7% Score 603; DB 27; Length 692;
Best Local Similarity 99.8%; Pred. No 0 000-00;
Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 88 GCTTGGCGTAATCATAGCTATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATTC 147
QY 4008 GCTTGGCGTAATCATAGCTATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATTC 3067

Db 148 CACATCAATATAGAGCGGAGAGCATAAAGGTGTGGGTGGCTTAATGAGTGAGCT 207
QY 3068 CACATCAATATAGAGCGGAGAGCATAAAGGTGTGGGTGGCTTAATGAGTGAGCT 3127

Db 208 AACTCAGCATTAATTCGTTGGGCTCATGTGGCGGCTTTTCAGTCCGSSAAAGCTGTGGTGCC 267
QY 4128 AACTCAGCATTAATTCGTTGGGCTCATGTGGCGGCTTTTCAGTCCGSSAAAGCTGTGGTGCC 3187

Db 268 AGCTGCATTAATGAATCGGCTAAAGCGGAGAGAGAGAGATTTTCGCTATTGGGCGCTCTT 327
QY 4186 AGCTGCATTAATGAATCGGCTAAAGCGGAGAGAGAGATTTTCGCTATTGGGCGCTCTT 3247

Db 428 AGGCTTCCTTCGCTTCATTCAGTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCG 387
QY 4248 AGGCTTCCTTCGCTTCATTCAGTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCG 3307

```

```

Db 388 CTCACCTCAAAAGGCGGTAATACAGGTTATCTTACAAATTCAGGGATATAAGCGGAAATCA 447
QY 3308 CTCACCTCAAAAGGCGGTAATACAGGTTATCTTACAAATTCAGGGATATAAGCGGAAATCA 4467

Db 448 TGTACGCAAAAAGGCGGCAATCAAAAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 507
QY 3368 TGTACGCAAAAAGGCGGCAATCAAAAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4477

Db 508 TCCATAGGCTCCGCGCCGCCCTGACGAGCATATCAAAAATTCGAGGCTCATATCAATAGGTCGC 567
QY 3428 TCCATAGGCTCCGCGCCGCCCTGACGAGCATATCAAAAATTCGAGGCTCATATCAATAGGTCGC 4487

Db 568 GAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
QY 3488 GAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4547

Db 628 CTTCTGTGTGCAAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 487
QY 3548 CTTCTGTGTGCAAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4607

Db 688 TGGCGG 692
QY 3608 TGGCGG 3612

```

Search completed: Tue Mar 16 12:04:59 1999
Job time : 8144 secs.

W P S R E H

(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nnn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Mar 16 00:33:53 1999; MasPar time 1504.62 Seconds
1341.307 Million cell updates/sec
Tabular output not generated.

Title: >US-09-056-767-1
Description: (1-854) from US09056767.seq
Perfect Score: 854
N.A. Sequence: 1 AAGCTTTGAAGATTGATTG... TATTTAAACACAAAGCTT 854
Comp: TTCGAACACTCTAAGCTAAC... ATAAATTTTGTGTTTCGAA

Scoring table: TABLE default
Gap 5

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 08
Listing first 45 summaries

Database:

emb156
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_ov 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_to 14:em_un 15:em_vi
genbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi

Statistics: Mean 10.619; Variance 4.803; scale 2.211

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	854	100.0	4998	34	PVYOCRNA Potato virus Y genome, 0.00e+00	
2	796	93.2	9704	34	PVYXXX Potato virus Y strain N 0.00e+00	
3	796	93.2	9704	34	PVYAAA Potato virus Y (N stra 0.00e+00	
4	796	93.2	9705	22	A08776 Potato virus Y complet 0.00e+00	
5	788	92.3	9694	34	PVU09509 Potato virus Y common 0.00e+00	
6	780	91.3	1308	22	I33135 Sequence 1 from patent 0.00e+00	
7	532	62.3	9703	34	PVYPOLYP Potato virus Y polypro 0.00e+00	
8	530	62.1	9701	34	PVYGEN Potato virus Y genes e 0.00e+00	
9	353	41.3	1784	34	PVYIGK2 Potato virus Y (strain 2.92e+290	
10	328	38.4	3258	34	PVU09508 Potato virus Y necroti 5.68e+267	
11	209	24.5	9640	34	PMVCG Pepper mottle virus co 4.20e+157	
12	122	14.3	9786	34	PPVCG Plum pox potyvirus com 8.31e+79	
13	120	14.1	9787	34	PPVCG Plum pox virus RNA gen 4.78e+77	

14	119	13.9	9741	34	PPV	Plum pox virus genomic	3.62e-76
15	118	13.8	1695	34	PPVNIAPRO	Plum pox potyvirus ORF	2.73e-75
16	106	12.4	4773	34	PPVPGP	Plum pox virus partial	8.24e-65
17	100	11.7	9786	34	PPVOPPRO	Plum pox potyvirus pol	1.29e-59
18	99	11.6	9496	34	TEVWNP	Tobacco etch virus Pl,	9.41e-59
19	91	10.7	726	22	I23338	Sequence 4 from patent	6.87e-52
20	91	10.7	4244	34	SPFC1	Sweet potato feathery	6.87e-52
21	91	10.7	9494	34	TEVCGHAT	Tobacco etch virus (hi	6.87e-52
22	91	10.7	9495	22	I31773	Sequence 1 from patent	6.87e-52
23	91	10.7	9497	34	TEVGEN	Tobacco etch virus RNA	6.87e-52
24	91	10.7	10820	34	D86371	Sweet potato feathery	6.87e-52
25	87	10.2	9585	34	PVCGA	Potato virus A complet	1.74e-46
26	84	9.8	4306	34	MBGB625	Bean yellow mosaic vir	6.04e-46
27	84	9.8	9532	34	MBGP	Bean yellow mosaic vir	6.04e-46
28	79	9.3	3684	34	OMVNIAB	Ornithogalum mosaic vi	9.70e-42
29	79	9.3	9475	34	TU038621	Tobacco vein mottling	9.70e-42
30	78	9.1	9860	34	PSMVGENOM	Pea seed-borne mosaic	6.56e-41
31	77	9.0	1320	10	E11059	DNA encoding clover ye	4.55e-40
32	77	9.0	3818	34	AB002698	Clover yellow vein pot	4.55e-40
33	77	9.0	9584	34	AB011819	Clover yellow vein vir	4.55e-40
34	70	8.2	729	34	TU018654	Turnip mosaic virus nu	2.86e-34
35	68	8.0	9475	34	POTTYVX	Tobacco vein mottling	1.25e-32
36	65	7.6	7218	22	I66494	Sequence 14 from paten	3.46e-30
37	63	7.4	10080	34	LMVOPP	Lettuce mosaic virus (1.43e-28
38	62	7.3	9547	34	BYU47033	Bean yellow mosaic pot	9.16e-28
39	62	7.3	9834	34	D83184	Turnip mosaic virus RN	9.16e-28
40	50	7.0	3525	34	RTMPARD	Shallot potyvirus (prc	3.67e-26
41	59	6.9	10086	34	PSU34972	Peanut stripe virus MR	2.30e-25
42	58	6.8	5102	34	TMVCAP	Turnip mosaic virus cd	1.44e-24
43	58	6.8	9830	34	TMVCAPP	Turnip mosaic virus ge	1.44e-24
44	56	6.6	2041	34	BCO012651	Bean common mosaic vir	5.47e-23
45	56	6.6	10080	34	LMVEPP	Lettuce mosaic virus (5.47e-23

ALIGNMENTS

RESULT 1
LOCUS PVYOCRNA 4998 bp ss-RNA VPL 01-SEP-1992
DEFINITION Potato virus Y genome, 3' half sequence.
ACCESSION D12539 D01241
NID Q22454
KEYWORDS RNA polymerase; coat protein; polypeptide; protease.
SOURCE Potato virus Y (strain 0) cDNA to genomic RNA.
ORGANISM Potato virus Y
REFERENCE 1 (sites)
Hidaka,M., Yoshida,Y., Masaki,H., Namba,S., Yamashita,S.,
Tsukizaki,T. and Uozumi,T.
Cloning and sequencing of the 3' half of a potato virus Y (O
strain) genome encoding the 5k protein, protease, polymerase and
coat protein
Nucleic Acids Res. 20 (13), 3515 (1992)
JOURNAL 92335011
MEDLINE 2 (bases 1 to 4998)
REFERENCE 2 (bases 1 to 4998)
Hidaka,M.
JOURNAL Unpublished (1992)
REFERENCE 3 (bases 1 to 4998)
Hidaka,M.
AUTHORS Direct Submission
TITLE Submitted (08-NOV-1991) to the DDBJ/EMBL/GenBank databases. Makoto
Hidaka, Faculty of Agriculture, The University of Tokyo, Department
of Biotechnology; 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
(E-mail:mamagoto@tansei.cc.u-tokyo.ac.jp,
Tel:03-3812-2111(ex.3082), Fax:03-5684-0387)
COMMENT Submitted (08-Nov-1991) to DDBJ by:
Makoto Hidaka
Department of Biotechnology
Faculty of Agriculture,
The University of Tokyo
1-1-1 Yayoi
Bunkyo-ku
Tokyo 113

VCQAIIIDQRYYREIELMQIEKNEIVCMELYSLOKRLERDFTWDEYIEYLKSVNPQ
 IVQFAQYERHSDYDVVRKNSPPVKNLEQVFAFMAVIMVFDABRSDCVWDLKQKGV
 LSSDQYVMEHOSLDVVKNFDERNEIIDFELSDTIRTSVLDTKESDWDKLOKQGH
 TLPHRYTEGHEPMETRATAQVANDIAHSEHDLFELRGVSGSKSLGVLPHLSVAGSV
 LLIEPFRFDECHVFKUSSEPPFKPILRMFGNSIPGSSPISVMTSGFALHYFANPFS
 QUAQNEFVIEACNVLDPSPSAMAFSLLSVYQACKVLKVSATPVGREVEFTTQOPVKL
 IVEDTSFQSFQAGSKTNADVYVQFNSVLYVSSYNEVDITLAKLLTDKNMWTYVD
 GRMKHGLEIVTGTSGARPHFVATNIENGVTLDIDVVDVFLGKVPFLDIDNRSI
 AYNKVSYSYGRIGORLGRVGRFKAVALRIGTEPKGIIIEIPSMVATEAALACAFYNLP
 MTPGDSIPRASSTWLSVSEYEPFLGVALEIPQVKIAFIHKEIPKLEHMLWETV
 KQVDFPSPIRASSISIAIYTRIDLFAPTRILVILLEERVKQSFRLIDEG
 CSMFVINTLNRARYAKDYTAENIKLEKRSQKLEFNSLDGSAEENLKRKVES
 LQFVHQAAATSLAKDLKLGKWNKSLVAKDIIAGAVAIIGGLIYSWFTQSVETVSH
 QGKNKXPIQALKPHAPDKRGAFEIDNDDTIEFFFGSAVRKKGKGGKTVGMKSS
 RRFNMVGFDPTEYSFIOQVDPDLTGRIEENVADIDIOERSEVRKMKVENDDIEM
 QALGNTIHAIFKDCWDKALKIDLPHPLKVCYDKTNGIAPFPERELEROTGPV
 EVDVDPQAEVEHAKSLMGLRDFNPQIAQVCLRLKVSVEYGASEMYGFGAYIVA
 ADKEIIVMSCLRLYKGLLGIWNGSLAKELCKEILANKTRITFAAPLDTLLGKVC
 VDDNNQYSKNIIECCVTGMKTEYGGWDLRLRRLPENWYVCADGSOFSLSPTVLI
 NAVLTIRSTYMEDVDGLQMLNLYTEIVTPISTPDGTIVKFRGNNSQSPVVDN
 SLVVLAHMYALIKCEVEFEIDSTCVFVNGDOLLJAVNPEKESIIDRMSQHSFSLG
 LNYDFSSTRKEBELWEMSHRGLLEGWYVPLKEERIVSLOWDRADLPEHRLAIC
 AAMIESGYSELTHQIRRYSWLQQQPFATIAQGRAPYIASMALRKLMDRAVDEE
 ELRAFTENMVALDEFELDSYEVHQAANDIDAGSNKKDAKPEQSIQPNPKKQDE
 DVNAGTSHTDVPRIKAITSKMRPTSKGATVPNLEHLEYAQQIDISNTRATQSF
 DTWEAVRMAYDIDETEMPTVMGLVMWCIENTGSPNVNVMVWMDNEOVEXPLKPI
 VENAKPTLQIRMAHFSDVAEAYIEMRNKKEPMRYGLIRLMDGLARYAFDFEVT
 SRTPVAREAHIQMAAKALKSAQPRLEGLDGGISTQENTERTHTEDVSFMSHTLGV
 KNM.

mat_peptide 185..1009
 /note="putative extreme 5'protein"
 old_sequence 788
 /citation="[1]
 /replace=""
 old_sequence 847
 /citation="[1]
 /replace="t"
 mat_peptide 1010..2656
 /note="putative"
 /product="helper component protein, HC"
 mat_peptide 2657..3655
 /note="putative"
 /product="38k protein"
 mat_peptide 3656..5557
 /product="cytoplasmic inclusion protein"
 mat_peptide 5558..5713
 /note="putative"
 /product="genome linked protein, Vpg"
 mat_peptide 5714..7009
 /note="putative"
 /product="nuclear inclusion protein, Nla"
 mat_peptide 7010..8572
 /note="putative"
 /product="nuclear inclusion protein, Nib"
 mat_peptide 8573..9373
 /note="putative"
 /product="capsid protein"
 BASE COUNT 3004 a 1818 c 2273 g 2609 t

Query Match 93.2%; Score 796; DB 34; Length 9704;

Best Local Similarity 96.8%; Pred. No. 0.00e+00;

Matches 825; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Db 6117 AACCTTTGAAGATTGATTAAATGCCACATAACCCACTCAAGCTTGTGACAAAACAATG 6176

|||||

QY 1 AACCTTTGAAGATTGATTGATGCCACATAACCCACTCAAAATTTGTGACAAAACAATG 60
 Db 6177 GCATTGCCAAATTTCTTGAGAGAGAGCTCGAACTIAGGGAGAGCTGGCCAGCTGTAGAAG 6236
 QY 61 GCATTGCCAAATTTCTTGAGAGAGAGCTTGAGCTIAGGGAGAGCTGGCCAGCTGTAGAAG 120
 Db 6237 TCGATGTGAAGGACATACACAGCAGCAGGAGGTGAGGATGAAGCTAAATCGCTCATGAGAG 6296
 QY 121 TCGAGTGAAGGACATACACAGCAGCAGGAGGTGAGGATGAAGCTAAATCGCTCATGAGAG 180
 Db 6297 GCTTGAGAGACTTCAACCCCAATTTGCCCAACAGTTTISAGGCTGAAGATATCTGTTGAAT 6356
 QY 181 GCTTGAGAGACTTCAACCCCAATTTGCCCAACAGTTTISAGGCTGAAGATATCTGTTGAAT 240
 Db 6357 ATGGGACATCAGAGATGTACGGTTTGGATTITGGAGCATACATAGTAGGAACCCACCAT 6416
 QY 241 ATGGGACATCAGAGATGTACGGTTTGGATTITGGAGCGTACATATAGTAGGAACCCACCAT 300
 Db 6417 TATTTAGGAGTTTCAATGGTTCATGGAGGTGCAATCCCATGCACGGTACATTCAGGGTGA 6476
 QY 301 TGTTCAGGAGTTTCAATGGTTCATGGAGGTGAGATCCCATGCACGGTACATTCAGGGTGA 360
 Db 6477 AGAATCTACAGCTTTGAGCGTTTCTGCCCAATTAAGGTAGGGACATCATCTCATCAAAA 6536
 QY 361 AGAATCTACAGCTTTGAGCGTTTCTGCCCAATTAAGGTAGGGACATCATCTCATCAAAA 420
 Db 6537 TGGCGAAGATTTCCCTGCTCTTCCACAGAAATTTGCAATTTCCGAGCTCTTACACAGAAG 6596
 QY 421 TGGCAAGATTTCCCTGCTCTTCCACAGAAATTTGCAATTTCCGAGCTCTTACACAGAAG 480
 Db 6597 AAAGAATTTCTTGTAGTTGAACCAACTTCCAAGAGAAGTATGCTTCGTGATCATCACAG 6656
 QY 481 AAAGAATTTCTTGTAGTTGAACCAACTTCCAAGAGAAGTATGCTTCGTGATCATCACAG 540
 Db 6657 AAACAAGCATCTTCAATATACAGGAGCAGACATCTTGAAGAGCATTTGGATTGAACAG 6716
 QY 541 AAGCAAGCATCTTCAATATACAGGAGCAGACATCTTGAAGAGCATTTGGATTGAACAG 600
 Db 6717 ATATGGACATTTGGACTACCGAGTGTGGAGCAGCCGAGTGTATAGTCGGAATTC 6776
 QY 601 ATATGGACATTTGGACTACCGAGTGTGGAGCAGCCGAGTGTATAGTCGGAATTC 660
 Db 6777 ACAGTCTGGCAAAATATGACACACACCACTACTACTACTAGCTTCGATGAAGATTITG 6836
 QY 661 ACAGTCTGGCAAAATATGACACACACCACTACTACTACTAGCTTCGATGAAGATTITG 720
 Db 6837 AAAGCAAGTACCTCCGACCAATGAGCACAATGAATGGGTCAAGCTCTGGGTTTAAATC 6896
 QY 721 AAAGCAAGTACCTCCGACCAATGAGCACAATGAATGGGTCAAGCTCTGGGTTTAAATC 780
 Db 6897 CAGACAGTGTGTGGGGCCCGTTGAAACTTAAAGACAGCAGCTCCGAAAGGTTATCA 6956
 QY 781 CAGACAGTGTGTGGGGCCCGTTGAAACTTAAAGACAGCAGCTCCGAAAGGTTATTA 840
 Db 6957 AAACAACAAGCTT 6970
 QY 841 AAACAACAAGCTT 854

RESULT 4
 LOCUS A08776 9705 bp RNA PAT 27-NOV-1993
 DEFINITION Potato virus Y complete genome RNA.
 ACCESSION A08776
 NID 9492885
 KEYWORDS
 SOURCE Potato virus Y.
 ORGANISM Potato virus Y.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 Potyvirus.
 REFERENCE 1 (bases 1 to 9705)
 AUTHORS
 TITLE COMPLETE NUCLEOTIDIC SEQUENCE OF THE COMPLEMENTARY DNA OF THE
 GENOMIC RNA OF THE POTYVIRUS, GENES CODING FOR THE POTYVIRUS

100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

TITLE Nucleotide sequence and genome organization of a Canadian iso-

Wed Mar 17 09:44:07 1999

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Can. J. Plant Pathol. 18, 209-214 (1996)
2 (bases 1 to 9698)
Singh, R.P.
Direct Submission
Submitted (10-MAY-1994) Rudra P. Singh, Fredericton Research
Centre, Agriculture and Agri-Food Canada, 850 Lincoln Road,
Fredericton, New Brunswick, E3B 4Z7, Canada
3 (bases 1 to 9698)
Singh, R.P.
Direct Submission
Submitted (25-MAR-1997) Rudra P. Singh, Fredericton Research
Centre, Agriculture and Agri-Food Canada, 850 Lincoln Road,
Fredericton, New Brunswick, E3B 4Z7, Canada
Sequence update by submitter
On Mar 25, 1997, this sequence version replaced gi:602403.

REMARK
COMMENT
FEATURES
source

1. .9698
Location/Qualifiers
/organism="Potato virus Y"
/strain="common strain"
/specific_host="Solanum tuberosum L."
/db_xref="taxon:12216"
/lab_host="Nicotiana tabacum L."
/clone="PO7"
/tissue-type="leaf"
125. .9370
/note="nuclear inclusion protein Nib and capsid protein"
/codon_start=1
/product="polyprotein"
/db_xref="PDB:1G90388"
/translation="METISTQGAISSLTPTNKKSSMATYMTICGFSPECKLPYSPASC
GLIVKEREVLASVFPADLEQLSARLLKQBYAYRVLKNGTETRYRYKQDAQIKRIQK
KLEDEPHEFYHFOAAPSIVSKITTAGGDPKSEPOAQLHHTPRVKYKTRIL
KLTEGQMHFIKQVIMSEKRSVHLNKKTHVQYKEILGAYSTVRAHMGUR
RVDFQDMVTGLLQRLARTKWSQVRIINRRGDSQVILNTRKSLKSGHSGDGLF
IVRSGHEGLIADSRVTSQVLSDMIQFENADNFWKGLDGNWARMYPSDHTCVAGLP
VDCGSAVMAHSLPCYKICTCAQOAYSLPVSDFLKLHKKHARDGNLRLGAKND
RTHVKNFLIAHLETPVDLNLLELFNEIKSIGEKQOAFKLNVLNNFLKCKENT
AHEWQAUSLELAFQAKNRDNIKKGDISFRNKLFAKANWNLISCDNDKKNAN
FLMGQRYHAKRFESNFEIDPAGKISAYELIRKPSGTRKLSIGNLVPLDIAERQ
KMGQDYKQVRSKTSKSDGNYVPCCTTLDGSAIESITFYPTPKKHLVTONSGD
QKFDVPLQDSEMLYIAQGYCIINFLAMHNTIGEDAKDTKQVMDMCPKLGTPW
TMDLATTCAQMEIFYDPAELPSLLVHDHTQCHVDVSFGSTGYHILKASSVS
QLILFANDELESDIKHYRVGVPNACPELGGTISFREGGVIMSESAALKLLKGFIR
PKYMRQLLDEPVLILSLSPCLLMANNNGIFELAVRLWIKQVAMIASLISAL
ALRVSALEDAQRIIDAAADLDDACDGNLHLYPTALWLVQVKNRNECDFTL
FKAGFPYSNTSVQIMEKNFLSLDDAWKDLWRKLSATWYSYRAKESITRYIKPTG
RADLGLYNSPQAFGRSAQVYKVTAGSLSERFNINIKCVNISSEFFIRIFRPLP
TFVTFVNSLLVISMVAVCAQIILDQRYRREIELMQIEKNEIVCMELYSIQRK
LERDFTWDEYIEVLKSVNPQIVQAOAMEEYDVRHSTFGVKNLEQVAFMALVIM
VEDAESDCVFKTLNFKVGLSLDHEVRHQSLDDVIKNFDERNEIIDFELSEDITRT
SSVLDTRFSDMORQIQMGHTLPHYRTEGHFMEFTATAVOVANDIAHSEHLDLVRG
VVDLTKLVDKNNMTVDGRTMKHGCLEIVTKGSARPHFVWATNIIENGVLID
VWDFGLKLVSPFLIDNRSIAYNKESYSGERIQRLRGVRFKGVALLRIGHTEKGI
EIPSWATEALACAFVAVMTGGVSTSLIGNCTVQVTKMQQFELSPFTIQNFHVA
DGSWHPVTHLLKYLKDCMTPLCQDSIPYRASCWLSVSEYERLVALLEIPQIKI
AFHKEIPPKLHEMLWETVYKVDCLFSPISIRASSISKIAYLTRDLFAIPTLIVE
RLLEERVQSQFSLIDEGCSNFSVNLNTLRLARYAKDYTAENIKLEKRSOLA
EFSNLDGSAEENLIKRYESLQVHQAATSLAKDLKLGKTKWSLVAKDLIIAGAVA
IGGILGYISFTQSVETVSHQGNKSKRIQALFRHARDKRAGFEDNNNDIYEFEG
SAYRKKGKGTQVGMGSKSRKFINMYFDPTIYSFIQVDPPLTGAEEENYADIRD
IQDRESEVRKLENDNDIEMOAGSNNTIHAVERKDSKALKIDLPNHPNLCVCKDT
NGIAFERRELELQTPAVENVVKDIPAOEVEHEAKSLMRGLRNPFAQVPCRLKY
SVEYGTSEWYGFQGAFCVIAHHNHLFRSNGSMVGHGTFRVKNLSLSVLPKGRD
IILIKMPDKFPVFPQKLFHFRAPIONERVCLVGNFOEKVASSIITETSTYINPQST
WKHLETIDNGHGLFPVSTADCLVGLHSLANNAHNTYSADEDFESKYLRAHEHN
EWKWSKNYNDPTVLWGLPKLKDSTPKGLFTKTLVDQLIEDHVDPMYQAKHSWMEAL
TGNLQAVATMSQLVTKHVKVGCGRHFEKFTLVDAAEAFRPMYDMSKSLNRLAD
IKDINMYKSIDVIGDCAFEAAINRVIYLVGHFKKAYVTDQEFKALNNKAA
VGMVGGKKDYFHEFTDADKEEIVMSCLRIYKGLLIGWNSLKAELRCRKKILANK
TRIFTAAPLDLTLGKGVCDVDFNNQYFSKNIECCVTGMYKEYGGWMDKLLRLPENWV

YCDADGSOSSLTPTLFLNAVLTIRSTYMEDWDVGLQMLRNLYTEIVVPISTPDGTI
VKFRGNSQSPSTVVDNSLMVYLAMHYAFIREGIEFEETISCTFFVNGDDLLIAVN
PKEDILDRLSQHSFSDLGLNYDFSRTNRKEELWFMSHRLIEGMYPKLEERIVS
LLOWRADLAERLEAICARMIESWGYSLTHQIRRFYSWLLQOQFPATIAEGKAPY
TASNALPKLYMDRADEEELRAFETMMVALDDDEFEDSEYVYHQANDIDAGGSNKKD
TKQSOSSIQSNPKGKDKDYNAGTSGTHTVPIKATSKMRMPKSKGAVALNLEHLE
YAPOQIDISNTRATQSFQEDTWYEAVPWYDGETEMPTVMNGMYWCIENGSPNVNG
VWMMGNEQVEYFKPELIVENAKPTLRQIMAHESDVABAYIEMRNKKPEYMPRYGLIR
NLRDMGLARYAFDEYVTSRTPVAREAQIOMRAAALKSNAQPLRFLDGGISTOENT
ERHTTEDVSPSMHTILLGVKNM"
182. .1012
/product="P1"
mat_peptide 1010. .2407
/product="HC"
mat_peptide 2405. .2659
/product="HC-pro"
mat_peptide 2657. .3502
/product="P3"
mat_peptide 3500. .3658
/product="6K1"
mat_peptide 3656. .5557
/product="C1"
mat_peptide 5558. .5716
/product="6K2"
mat_peptide 5714. .6277
/product="N1a-Vpg"
mat_peptide 6278. .7009
/product="N1a-Pro"
mat_peptide 7010. .8566
/product="N1b"
mat_peptide 8567. .9367
/product="CP"
BASE COUNT 2995 a 1818 c 2273 g 2612 t
ORIGIN
Query Match 92.3%, Score 788, DB 34, Length 9698:
Best Local Similarity 96.1%; Pred. No. 0.00e+00;
Matches 821; Conservative 0; Mismatches 33; Indels 0; Gaps 0:
Db 6117 AAGCTTTGAAGATTGACTTAATGCCACACACACCCACTCAAGTTTGTGCACAAAACAATG 6176
QY 1 AAGCTTTGAAGATTGACTTAATGCCACACACACCCACTCAAGTTTGTGCACAAAACAATG 50
Db 6177 GCAATTCGAAAATTTCTCTGAGAGAGCTCGCACTAAGCAGAGCTGGCCAGCTGTAGAAG 6236
QY 61 GCAATTCGAAAATTTCTCTGAGAGAGCTCGCACTAAGCAGAGCTGGCCAGCTGTAGAAG 120
Db 6237 TCAATGTGAAGGACATACACGACAGAGGTGGAGCATGAAGCTAAATCGCTCATGAGAG 6296
QY 121 TCAATGTGAAGGACATACACGACAGAGGTGGAGCATGAAGCTAAATCGCTCATGAGAG 180
Db 6297 GCTTTGAGAGACTTCAATCCAAATGCCCCAAACAGTTTGTAGCTGAAAGTATCTGTTCAAT 6356
QY 181 GCTTTGAGAGACTTCAATCCAAATGCCCCAAACAGTTTGTAGCTGAAAGTATCTGTTCAAT 240
Db 6357 ATGGGACATCAGAGATGTACGGTTTGGATTGGAGCATACATATAATACGGAACCAACCAAT 6416
QY 241 ATGGGACATCAGAGATGTACGGTTTGGATTGGAGCATACATATAATACGGAACCAACCAAT 300
Db 6417 TGTTTAGGAGTTACAAATGGTTCCATCGAGGTGCAATTCATCGAGGTACATTCAGGGTGA 6476
QY 301 TGTTTAGGAGTTACAAATGGTTCCATCGAGGTGCAATTCATCGAGGTACATTCAGGGTGA 360
Db 6477 AGAATCTACACACTTTGACGGTTTGCCTTAAAGGTAGGAGCATATCTCTCATCAAAA 6536
QY 361 AGAATCTACACACTTTGACGGTTTGCCTTAAAGGTAGGAGCATATCTCTCATCAAAA 420
Db 6537 TGCCGAAAGATTTCCCTGTCTTTTCCACAGAAATTCGATTTCCGAGCTCCTATACAGATG 6596
QY 421 TGCCGAAAGATTTCCCTGTCTTTTCCACAGAAATTCGATTTCCGAGCTCCTATACAGATG 480
Db 6597 AAAGAGTTTGTAGTTGGGACCAACTTTTCAGGAGAGTATGCATCGTCAATCAACAG 6656

gene 1015..2409
 /gene="HC"
 mat_peptide 1015..2409
 /gene="HC"
 gene 2410..3504
 /product="helper component protein"
 /gene="p3"
 mat_peptide 2410..3504
 /gene="p3"
 /product="p3 protein"
 gene 3505..3660
 /gene="6K1"
 mat_peptide 3505..3660
 /gene="6K1"
 /product="6K1 protein"
 gene 3661..5562
 /gene="ci"
 mat_peptide 3661..5562
 /product="ci protein"
 gene 5563..5718
 /gene="ci"
 mat_peptide 5563..5718
 /product="ci protein"
 gene 5719..6282
 /gene="6K2"
 mat_peptide 5719..6282
 /product="6K2 protein"
 gene 5719..6282
 /gene="vpq"
 mat_peptide 5719..6282
 /product="vpq protein"
 gene 6283..7014
 /gene="Nia-Pro"
 mat_peptide 6283..7014
 /product="Nia-Pro"
 gene 7015..8571
 /gene="Nia-Pro"
 mat_peptide 7015..8571
 /product="Nia-Pro"
 gene 8572..9372
 /gene="Nlb"
 mat_peptide 8572..9372
 /product="Nlb protease"
 gene 9376..9701
 /gene="cp"
 mat_peptide 9376..9701
 /product="coat protein"
 3'UTR 9701
 polyA_site 9701
 BASE_COUNT 3057 a 1755 c 2238 g 2651 t
 ORIGIN

Query Match 62.1%; Score 530; DB 34; Length 9701;
 Best Local Similarity 81.1%; Pred. No. 0.00e+00;
 Matches 691; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

Db 5124 GCTCTAAAATGATTTGATGCCACACCAACCCACTCAAAATCTGTGATAAATCGCATTCGC 6183
 QY 3 GCTTTCAGATTTGATTTGATGCCACATCAACCCACTCAAAATTTGTGACAAACAAATGCG 62
 Db 5184 ATTGCTAAGTTTCCTGAAGAGAACTTCAGTTGAGGCAAACTGGCCAGCAATAGAGGTT 6243
 QY 53 ATTGCCAATTTCTGAGAGAGAGTTGAGTGAAGGAGAGTGGCCAGCTGTAGAGTC 122
 Db 6244 GATGGAAGACATTCACAAACAGGAAGTGGAGCATGAAGCAATCACTCATCAGAGGT 6303
 QY 123 SAGGTGAAGGACATACCAGCAGGAGGTGGAACATGAAGCTAAATCGCTCATCAGAGGC 182
 Db 6304 TTAAGGATTTCAATCCCAATTCCTCAACAGTTTCAGAGTAAAGTGTCTGTTGAATAT 6363
 QY 183 TTGAGAGACTTCAACCAATTCGCAACAGTTTGTAGGCTGAAGATATCTGTTGAATAT 242
 Db 6364 GCAAGTCTGAAATATGATGGTTGGTGTATATATATATATATATATATATATATATATAT 6423
 QY 243 GGGACATCAGAGATGACGGTTTGGATTTGGAGCGTACATATAGGCAACCCACCATTTG 302

Db 6424 TTCAAGAGCTTCAATGGATCCATGGAAGTGGCATCAATGCATGGACATTCAGAGTGAAG 6483
 QY 303 TTCAGAGATTATAATGGTTTCCATGGAGGTGATCCATCCAGCGGTACATTCAGGTTAAAG 362
 Db 6484 AATTTGCATAGCTTGAGGCTTTTACCGATCAAGGCAGAGACATTCATCATATAAAGATG 6543
 QY 363 AATCTACACAGTTTGAGCGTTCTGCCAATTAAGGTAGGACATCATCTCATTAATG 422
 Db 6544 CCAAGAGATTTCCCTGTTTTCACCAAAAACATGACCTTCCAGAGTCCAGTCCAGAAATGAG 6603
 QY 423 CCAAGAGATTTCCCTGTTTTCACCAAAAACATGACCTTCCAGAGTCCAGTCCAGAAATGAG 482
 Db 6604 AGGATTTGTTGGTGGAACTAATTTTCAAGAAAAACATGCATCATCAATCATCACAGAA 6663
 QY 483 AGAATTTGTTAGTTGGAAACCACTTTTCAAGAGAGTATGTCATCGCATCATCACAGAA 542
 Db 6664 ACGAGTACTACATCAATGTACCGGGCAGCAGCTTTTGGAGCATTTGGATTGAACAAAT 6723
 QY 543 GCAAGCACTACTTACAATATACAGCAGCAGCAGCATCTTGAAGCATTTGGATTGAACAGAT 602
 Db 6724 GATGGCATTTGGATTACCACTAGTAGTACAGCTGATGATGATGATGATGATGATGATGAT 6783
 QY 603 AATGGACACTGTGGACTACAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 662
 Db 6784 AGCTTGGCAATTAATGTGCAAAACACGAAATTTATTTTATGAGCTTTTATGAGGATTTTGA 6843
 QY 663 AGTTTGGCAACAAATGCACACACAGCACTACTACTAGCTTTCGATGAGATTTTGA 722
 Db 6844 AGTAAGTATCTCCGAACCTAATGAGCATATAGTGTGAGCAAAATCGTGGTATATTAACCA 6903
 QY 723 AGCAAGTACTCCGAACCAATGAGCACAATGAATGGTCAAGTCTTGGATTTTAATCCA 782
 Db 6904 GATAGTGTGTTGGGTCCTTGAAGCTCAAGGAGAGTACCCCTTAAGGCTGTTTAAAG 6963
 QY 783 GACACAGTGTGTTGGGCTTGAAGCTTAAAGAGACGACCTCTTAAAGGCTTATTTAA 842
 Db 6964 ACAACAAAACCTT 6975
 QY 843 ACAACAAAACCTT 854
 RESULT 9
 LOCUS PVYCI6K2 1784 bp DNA VRL 24-JAN-1994
 DEFINITION Potato virus Y (strain O) CI, 6K2, VPg, NIa-Pro, NIa gene.
 ACCESSION Z29526
 NID g450360
 KEYWORDS 6 kd protein; 6K2 gene; ci gene; Cytoplasmic inclusion protein;
 genome-linked protein; ni gene; polyprotein; protease; Sma.I
 nuclear inclusion protein; VPg gene.
 SOURCE Potato virus Y.
 ORGANISM Potatoes.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 Potyvirus.
 REFERENCE 1 (bases 1 to 1784)
 AUTHORS Weinicki, M.A. and Baulcombe, D.C.
 TITLE The nucleotide sequence of the central fragment of potato virus Y
 (O strain) genome encoding 6K2 and VPg proteins
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1784)
 AUTHORS Weinicki, M.A.
 TITLE Direct Submission
 JOURNAL Submitted (19-JAN-1994) Weinicki M. A., Institute of Biochemistry
 and Biophysics, Department of Protein Biosynthesis, ul. Pawinskiego
 5A, Warszawa, Poland, 02-106
 FEATURES
 source
 1..1784
 /organism="Potato virus Y"
 /strain="O"
 /isolate="British"
 /db_xref="taxon:12216"
 /clone="pPVY15"
 1..850
 gene

3' UTR	9375	.9640
BASE COUNT	3051 a	1747 c
	2235 a	2607 t

[illegible]

QY	567	GGCAGC	CAATTC	TCTGGA	AGCATG	GATTGAA	ACAGAT	AATGG	CACAC	TGTGG	ACTAC	CAAG	TG	625
Db	6759	GTGAGC	ACTCAG	ATGCG	AGATTC	TTCTT	GGGCT	CACAG	CTTGC	CAATTC	CAAG	CAAC	6818	
QY	627	GTGAGC	ACTCGG	ATGGAT	GCTAG	TGCG	GAATTC	CACAG	TTTGG	CAACA	CAATG	CACAC	686	
Db	6819	CAGA	TTTCT	TATG	CAGG	TTTCC	CTGACA	CAATTC	GAGAC	CACTAC	TTGT	CAAA	6878	
QY	687	ACGAC	TACTAC	TACG	CTTC	GATGA	AGATTT	GAA	GCAAG	TACCT	CCGAC	CAATG	746	
Db	6879	AATG	ATACT	GGATA	AGCAG	TGGCG	GTAC	AATC	CGGAT	GAA	GTTC	TGCGG	6931	
QY	747	CACA	TGAAT	GGGTCA	AGTCT	TGGAT	TTAAT	CCGAC	GACAC	AGTGT	TGTGGG	799		

RESULT	15	PPVNIAPRO	1695 bp	mRNA	VR1	03-AFP-1092
LOCUS		plum pox polyvirus	ORF containing NI-a-like protein mRNA, partial cds.			
DEFINITION						
ACCESSION		M26965				
NID		G333303				
KEYWORDS		NI-a-like protein; cleavage protein.				
SOURCE		Plum pox polyvirus (strain Rankovic)	cdNA to mRNA.			
ORGANISM		Plum pox virus				
		Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae: Potyvirus.				
REFERENCE		1 (bases 1 to 1695)				
AUTHORS		Garcta,J.A., Riechmann,J.L. and Lain,S.				
TITLE		Proteolytic activity of the plum pox polyvirus NI-a-like protein in <i>Escherichia coli</i>				
JOURNAL		Virology 170, 362-369 (1989)				
MEDLINE		89268456				
FEATURES		Location/Qualifiers				
		1..1695				
		/organism="Plum pox virus"				
		/strain="Rankovic"				
		/db_xref="taxon:12211"				
		1..1695				
		1				
		CDS				

[illegible]

Release 3.1A John F. Collins, BioComputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mitsch_on n.a. n.a. database search, using Smith-Waterman algorithm
Run on: Sun Mar 14 05:27:25 1999. Maspar time 5165.00 seconds
Tabular output not generated.
1538 600 Million cell updates/sec
Title: SUS-09-020-716-1
Description: (12363) from US09020716 seq
Perfect Score: 1463
N.A. Sequence: 1 TCGACTTCACGGGAGGAGC... GCGGAAACACACAGGTACC
Comp: ACGTGGAGCTCCGCGGAGG...
Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0: Query 0
Searched: 602357 seqs. 1181590623 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_lo 8:em_lo 9:em_lo 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vi
Database: qenbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_lo 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi
Statistics: Mean 11.929; Variance 5.483; scale 2.176
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result		Query		Match		Length		DB		ID		Description		Pred. No.	
No.	Score	No.	Score	No.	Score	No.	Score	No.	Score	No.	Score	No.	Score	No.	Score
1	2907	95.4	4432	32	EV00112N	Expression vector pC0112N	0.00e+00	1	2907	95.4	4432	32	EV00112N	Expression vector pC0112N	0.00e+00
2	2897	86.1	5641	32	AF016404	Chloroplast transformant	0.00e+00	2	2897	86.1	5641	32	AF016404	Chloroplast transformant	0.00e+00
3	2883	85.7	4433	32	EV00114N	Expression vector pC0114N	0.00e+00	3	2883	85.7	4433	32	EV00114N	Expression vector pC0114N	0.00e+00
4	2978	85.6	2057	32	U02449	Cloning vector pDIRECT	0.00e+00	4	2978	85.6	2057	32	U02449	Cloning vector pDIRECT	0.00e+00
5	2951	85.2	4133	32	U01568	Phagemid cloning vector	0.00e+00	5	2951	85.2	4133	32	U01568	Phagemid cloning vector	0.00e+00
6	2786	82.8	6824	32	A25909	Yeast integration vector	0.00e+00	6	2786	82.8	6824	32	A25909	Yeast integration vector	0.00e+00
7	2737	81.4	6160	16	EC17DN3	E. coli pT7HCH pl. DNA	0.00e+00	7	2737	81.4	6160	16	EC17DN3	E. coli pT7HCH pl. DNA	0.00e+00
8	2737	81.4	8953	16	EC17DN2	E. coli pT7HCH pl. DNA	0.00e+00	8	2737	81.4	8953	16	EC17DN2	E. coli pT7HCH pl. DNA	0.00e+00
9	2737	81.4	11271	16	EC17DN1	E. coli (HM174) pT7HCH	0.00e+00	9	2737	81.4	11271	16	EC17DN1	E. coli (HM174) pT7HCH	0.00e+00
10	2665	79.2	2863	32	AF013597	Phagemid cloning vector	0.00e+00	10	2665	79.2	2863	32	AF013597	Phagemid cloning vector	0.00e+00
11	2612	77.7	4854	32	AS14584	Artificial sequence; v	0.00e+00	11	2612	77.7	4854	32	AS14584	Artificial sequence; v	0.00e+00
12	2557	76.0	5350	32	AB003468	Cloning vector pAP3neo	0.00e+00	12	2557	76.0	5350	32	AB003468	Cloning vector pAP3neo	0.00e+00
13	2535	75.4	2863	32	SYNPT7190	Cloning vector pPhagem1	0.00e+00	13	2535	75.4	2863	32	SYNPT7190	Cloning vector pPhagem1	0.00e+00

SUMMARIES

14 2534 75.3 2860 32 SYNPT7190 Cloning vector pPhagem1 0.00e+00
15 2529 75.2 2927 10 E11370 Nucleotide sequence of 0.00e+00
16 2529 75.2 2927 10 E11370 Nucleotide sequence of 0.00e+00
17 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
18 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
19 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
20 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
21 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
22 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
23 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
24 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
25 2513 74.7 2828 32 CVFSE16E Cloning vector pPS16A 0.00e+00
26 2506 74.5 2887 32 SYNPT125V pPS16E Cloning vector 0.00e+00
27 2507 74.5 2887 32 SYNPT125V pPS16E Cloning vector 0.00e+00
28 2507 74.5 2887 32 SYNPT125V pPS16E Cloning vector 0.00e+00
29 2506 74.5 2828 32 SYNPT125V pPS16E Cloning vector 0.00e+00
30 2499 74.3 2919 32 XXU13867 pPS16E Cloning vector 0.00e+00
31 2494 74.2 2893 32 XXU13867 pPS16E Cloning vector 0.00e+00
32 2494 74.2 2893 32 XXU13867 pPS16E Cloning vector 0.00e+00
33 2494 74.2 2893 32 XXU13867 pPS16E Cloning vector 0.00e+00
34 2449 72.8 4234 32 AB007532 Signal sequence detect 0.00e+00
35 2439 72.5 3033 32 CVGEM7190 Ligation-independent c 0.00e+00
36 2439 72.5 3033 32 CVGEM7190 Ligation-independent c 0.00e+00
37 2439 72.5 3033 32 CVGEM7190 Ligation-independent c 0.00e+00
38 2431 72.3 4109 32 FSPKAT1 Cloning vector pPS16A 0.00e+00
39 2419 71.9 4486 32 EVGEM116N Expression vector pPS16A 0.00e+00
40 2411 71.7 3003 32 CVGEM72FF Cloning vector pPS16A 0.00e+00
41 2411 71.7 3003 32 CVGEM72FF Cloning vector pPS16A 0.00e+00
42 2401 71.4 3639 32 SYNPKA1 Cloning vector pPS16A 0.00e+00
43 2395 71.2 5657 32 AF053733 Expression vector pPK1 0.00e+00
44 2395 71.2 5657 32 AF053733 Expression vector pPK1 0.00e+00
45 2383 70.9 3457 16 ETPP4RIA E. coli plasmid pPS4 PN 0.00e+00

ALIGNMENTS

1 EV00112N 4412 bp DNA SYN 18-MAR-1998
LOCUS Expression vector pC0112N (modified from pC0112 in [2]).
DEFINITION 229587
ACCESSION 229587
NID g452344
KEYWORDS ampicillin resistance, beta-lactamase, ColEI origin, expression
vector, nos terminator, phase II region, rice actin1 promoter.
SOURCE expression vectors.
ORGANISM expression vectors.
REFERENCE 1 (bases 1 to 4412)
AUTHORS Liu, L., Dasgupta, I., Davies, J.W. and Hall, R.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4412)
AUTHORS McElroy, D., Blowers, A.D., Jones, B. and Wu, R.
TITLE Construction of expression vectors based on the rice actin 1 (Act1)
5' region for use in monocot transformation
J. Mol. Appl. Genet. 1 (6), 561-573 (1992)
JOURNAL Mol. Gen. Genet. 231 (1), 150-160 (1991)
MEDLINE 92092956
REMARK (sites)
REFERENCE 3 (bases 27 to 282)
AUTHORS Taylor, A., Stachurski, S., Phares, P., Jantzen, J. and Joehman, H.W.
TITLE Nopaline synthase, transcript mapping and DNA sequence
J. Mol. Appl. Genet. 1 (6), 561-573 (1992)
REFERENCE 4 (bases 1 to 4412)
AUTHORS Liu, L.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1994) Liu, L., John Innes Institute, Virus
Research, Colney Lane, Norwich, United Kingdom, NR4 7UH
FEATURES
source
1..4412
/organism="expression vectors"
/db_xref="taxon:33776"
misc_feature 1..30
/note="multiple cloning sites"

[illegible][illegible]

DB 1024: TCAGAGTCGACGACCGGTAGAGAAATCAACAAAGCAGTCTCTTTSACATGCTTTTCTTCTGCGC 10290
 CP 1106: GAGAGTATACAGCTGATACAAAGATACAAAGATCTCTTTSACATGCTTTTCTTCTGCGC 1047
 DB 10291: HAAATCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10350
 CP 1046: HAAATCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10350
 DB 10451: CAAGAGTACCAATCTTTTCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10410
 CP 986: CAAGAGTACCAATCTTTTCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 927
 DB 10411: AATGCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10470
 CP 926: AATGCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 867
 DB 10471: ACATACCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10530
 CP 866: ACATACCTGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 807
 DB 10531: GTTACGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10590
 CP 747: GTTACGCTTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 747
 DB 10551: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10550
 CP 746: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 687
 DB 10651: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10710
 CP 627: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 627
 DB 10711: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10770
 CP 626: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 626
 DB 10771: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10830
 CP 597: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 597
 DB 10831: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10890
 CP 447: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 447
 DB 10891: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 10950
 CP 387: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 387
 DB 10951: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11010
 CP 327: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 327
 DB 11011: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11070
 CP 267: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 267
 DB 11071: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11130
 CP 266: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 266
 DB 11131: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11190
 CP 147: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 147
 DB 11191: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11250
 CP 87: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 87
 DB 11251: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 11271
 CP 66: GAGAGTCTGCAAA'AAAAAAGCAGCGCTACGACGGTGGTTTGGTTTSCGGGAT 66

RESULT 10
 LOCUS AF013597 2863 bp DNA circular SYN 03 DEC 1997
 DEFINITION Phagemid cloning vector pATCH1, complete sequence.
 ACCESSION AF013597
 NID 92654777
 KEYWORDS phagemid cloning vector pATCH1,
 phagemid cloning vector pATCH1,
 artificial sequence; cloning vectors.
 ORGANISM
 REFERENCE 1 (bases 1 to 2863)
 AUTHORS Kast, P.
 TITLE pKSS--a second-generation general purpose cloning vector for
 efficient positive selection of recombinant clones
 JOURNAL Gene 138 (1-2): 109-114 (1994)
 MEDLINE 94171019
 REFERENCE 2 (bases 1 to 2863)
 AUTHORS MacBeath, G. and Kast, P.
 TITLE Read-through artifacts -- when popular gene expression systems need
 a PATCH
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 2863)
 AUTHORS Kast, P. and MacBeath, G.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-1997) Department of Molecular Biology and
 Chemistry, The Scripps Research Institute, 10550 North Torrey Pines
 Road, La Jolla, CA 92037, USA
 FFATYPES
 source 1. 2863
 /organism="phagemid cloning vector pATCH1"
 /db_xref="taxon:64188"
 /join(1..18,449..2863)
 /note="derived from Genbank Accession Number U01668"
 /citation(1)
 19..77
 /note="derived from synthetic plasmid template pKSS"
 /citation(2)
 22..48
 /note="six-histidine tag for affinity purification of
 fusion proteins, and collected stop signals"
 76..438
 /note="derived from vector pET 23b(+)" (Novagen, Madison,
 WI)
 206..661
 /note="f1 phage origin region for single strand DNA
 replication; direction=right"
 498-499
 /note="nicking site of gene 11 protein"
 792..1652
 /gene="bla"
 792..860
 /gene="bla"
 792..1652
 /gene="bla"
 /function="ampicillin resistance"
 /codon_start(1)
 /transl_table(1)
 /product="pre-beta-lactamase"
 /db_xref="pid:a2654778"
 /translation="95-IGSRKVALLEFLAALCLLEVAIRLELVKALHAGQVAPVTH
 IELDLSKRLDSFPEERFEMSTFKVLGGAVLSRIADGQDGRPHYSNDVAF
 YSFVTEFLDTGCTVEELCSAATMSNTAAALLETTSPTPEELATFHRSDHVP
 DWDFEFLFATIDNEDPTTDMVAVATLPLKLTGLPLTLASPTGLIHWMLAKVAGLE
 LPSALPACWFLANWSCAGPEPSGCTLAALTNWTFSTVAVVATLSTGALMLPQPLGA
 EIGASLLKHW"
 1912..2426
 /note="pMB1 (ColE1-like) plasmid origin of replication; as
 in pBR322; the DNA region indicated is sufficient for a
 functional replication origin; direction=right"
 2411..2413
 /note="DNA replication initiation site"
 <2594..2686
 /gene="lacI"

Cp 500 TAAACACGCTGCGGCGGAGTACCAAGACGAGCGAGCGAGTACGAGCGAGCGAGCGGGA 301
 Db 6931 aaagagcgaataacgaacacgcctctcccccagcttgaccgaattcattatcgaagctg 6992
 Cp 500 AAGAGCGGCAATAGCGAAAGCGCTGCGGCGGAGTACCAAGACGAGCGAGCGAGCGGGA 241
 Db 6993 gacacacaggttcccgactgaagcgaggcgagtgagcgcaacgaataatgtgagtt 7052
 Cp 240 GACGACAGCGTTCGCAATGCAAGCGAGCGAGTACCAAGACGAGCGAGCGAGCGGGA 181
 Db 7053 gctcactcattaggaacacccagcgtttacactttatctctccgctcgtatgtgtg 7112
 Cp 180 GATCAATCAATAGCGAGCGAGTACCAAGCGAGTACCAAGCGAGCGAGCGGGA 121
 Db 7113 aattgtgagcgatatacaatttcacaggaagaaacagctatgacatgattacgcaagct 7172
 Cp 120 AATTGTGAGCGGATACCAATTCACAGGAGCAACAGAGTATGACCATGATTACGCGCAAGCG 61
 Db 7173 ggaagatgaacgtcactgaagggggaacaaagctg 7207
 Cp 60 CGCAATTAACGCTCACTAAGCGGAGCAACAAAGCTGG 26

RESULT 3

ID 507418 standard: DNA: 5585 BP
 AC 087418:
 DT 21-OCT-1998 (first entry)
 DE Plasmid pMPE-IL2.
 KW AAV: liposome; gene transfer; vector; adeno-associated virus;
 KW gene therapy; interleukin-2; tumor infiltrating lymphocyte; HIV.
 OS Synthetic.
 FH Key
 FI misc_rna
 FI 1..44
 FI /*tag= a
 FI /*note= "Bluescript KSII+ backbone"
 FI 45..239
 FI misc_rna
 FI /*tag= b
 FI /*note= "AAV left terminal region"
 FI 243..1076
 FI promoter
 FI /*tag= c
 FI /*label= CMV_promoter
 FI 1079..1264
 FI misc_rna
 FI /*tag= d
 FI /*note= "adeno virus major late intervening
 FI sequence"
 FI 1269..1357
 FI misc_rna
 FI /*tag= e
 FI /*note= "mouse Ig intervening sequence"
 FI 1394..1564
 FI 5'utr
 FI /*tag= f
 FI /*note= "rat preproinsulin 5' UTR"
 FI signal_peptide
 FI 1565..1579
 FI /*tag= g
 FI /*note= "rat insulin signal peptide"
 FI 1583..2071
 FI cds
 FI /*tag= h
 FI /*product= human IL-2
 FI 2085..2471
 FI poly_a_signal
 FI /*tag= i
 FI /*note= "SV40 poly-A signal"
 FI 2579..2762
 FI misc_rna
 FI /*tag= j
 FI /*note= "AAV right terminal region"
 FI 2763..5585
 FI misc_rna
 FI /*tag= k
 FI /*note= "Bluescript KSII+ backbone"
 PN W09507995-A.
 PD 23-MAR-1995
 PE 13-SEP-1994; U09774.
 PR 13-SEP-1993; US-120605.
 PA (IMMU-) APPLIED IMMUNE SCI INC.
 PI Lebkowski J, Philip R.

DP WPI: 95-131363/17.
 PT Liposome(s) contg. adeno-associated viral material, esp. plasmid
 PT - useful for introducing genes in lymphocytes, tumour cells etc.,
 PT e.g. for treating cancer, HIV infection, autoimmune disease or
 PT genetic abnormalities, also isolated T effector cells.
 PS Disclosure: Fig. 3A-3G; 92pp; English.
 CC Plasmid pMP-IL2 comprises a human IL-2 expression cassette flanked
 CC by AAV sequences. The vector is useful for gene transfer (e.g.
 CC by lipofection) into T-cells, tumor infiltrating lymphocytes and
 CC neoplastic cells from solid or fluid sources as a means of gene
 CC therapy, tumor vaccination, etc.
 SQ Sequence 5585 BP; 1394 A; 1401 C; 1399 G; 1391 T;
 Query Match 83.9%, Score 2820; DS 14; Length 5585;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 2821; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 2764 agatccaattcgccctatagtgactctattacgcgcgcctcactgcgcctctttacaa 2823
 Cp 2884 AGCTCCAAATCGCCCTATAGTGAUTCGTATACGCGGCGTCAATGCTCTCTGCTTTACAA 2825
 Db 2824 gctcgtactgggaacccctggcggtttaccacacttaacgccttgcagacacacccct 2883
 Cp 2824 CGTCGTGACTGCGAAACCCCTGGCTTTACCGCACTTAATGCGCTTGGAGACATCGCCCT 2765
 Db 2884 ttgcgcagctggcgtatagcgaagagggccgcacccatcgcccttcccaacagattacg 2943
 Cp 2764 TTCTGCTACGCTCGGTNA:AGCGAGAGAGCGCGCGTACGCGTATGCGCGTATGCGCGT 2705
 Db 2944 agcctgaattggcgaaatggcgccctgtagcgggcgcatgaagcgcggggtgtgtg 3003
 Cp 2704 AGCCTGAA*GGCGAATGGGACGCGCGCTGTAAGCGGCGCTATTAAGCGGAGGCGTGTG 2645
 Db 3004 gttacgcgcagctgaacgcgtacacttccacgcgccttagcgcgcctctttcgccttc 3063
 Cp 2644 GTTACGCGCAGCTGACCGCTACACTTCTGAGCGCTTAGCGCGCGCTCTTTTGGCTTTC 2585
 Db 3064 ttcccttctcttgcgcacgcttcgcgcgctttcccccgcacgctcgaagctcgaatcggggactc 3123
 Cp 2584 TTCCCTCTCTTTCTGCGCAGCTTGGCGGCTTTCCCGCGCTCAAGCTCAATATGGGGGCTC 2525
 Db 3124 ccttagggttcgagtttagctcttaaggacactcgaccccaaaaacttgattaggt 3183
 Cp 2524 CTTTAGGGTTCGATTATAGTCTTTAGCGGACCTGACCGCGCGCGCGCGCGCGCGCGCG 2465
 Db 3184 gatggttcacgtagtgggcctgcgcctgatacagcggtttttgcgcctttcactttgag 3243
 Cp 2464 GATGTTTACGTAGTGGGCTATGCGCGCTGATAGACGCGTCTTCTGCGCTTGGAG 2405
 Db 3244 tccacgtttttaaataagtgagctctgttccaaaactggaacacacactcaacccatctcg 3303
 Cp 2404 TCGACGTTCTTTAATAATGAGCTCTTGTCTCAAACTGGAACACACTCAACCTCATCTCG 2345
 Db 3304 gctattcttttgattataagggattttgcgcgattttgcgcctattgttttaaaatgaag 3363
 Cp 2344 GTCTATCTCTTTGATTATAGCGGATTTTGGCGGATTTAGCGGCTATTTGTTAAAAAATGAG 2285
 Db 3364 ctgatttaacaaaatttaacgcgaattttaacaaaatttaacgcgaatttaacgcgaatttaggtg 3423
 Cp 2284 CTGATTTACAAAAATTTAAAGCGAAATTTAAACAAAATTAACGCTTACAAATTTAGGCTG 2225
 Db 3424 gcacttttcgggaaatgtgcgcggaacccctattgttttcttaaaatacattcaa 3483
 Cp 2224 GCATTTTTCGGGGAATTTTGGGAGAACTTATTTGTTTATTTTCTAAATACATTCAA 2165
 Db 3484 atagtatccgctcatgagacaaataaacctcgatataatgcttcaataatttgaagaaga 3543
 Cp 2164 ATATGATCCGCTCATGAGACAAATAAGCGTATAAAATGCTTCAATTAATTAATGAAAGGA 2105
 Db 3544 agagtatgattatcaacatttcgctgctgcgccttattcccttttttgcgcattttacc 3603
 Cp 2104 AGAGTATGAGTATTCACATTTCCGCTGCGCTTATTCGCTTATTCGCGGCAATTTTGGC 2045

Db 920 gctatggttaaaatgagctgatttaacaaaatttaacgcaatttttaacaaaata 979
 Cp 2305 gctatggttaaaatgagctgatttaacaaaatttaacgcaatttttaacaaaata 2246
 Db 980 ttaacgtttacuatcttcaagtgagcattttcgaggaaatgtcccggaacccctattgt 1039
 Cp 2345 ttaacgtttacuatcttcaagtgagcattttcgaggaaatgtcccggaacccctattgt 2187
 Db 1040 ttaattttcaatatacatcaaatatgtatccctcatagagacaataacccctgataaatg 1099
 Cp 2184 ttaattttcaatatacatcaaatatgtatccctcatagagacaataacccctgataaatg 2127
 Db 1100 cttaataatattgaaagaaaggaagggatagatattcaacatttcggtgcacctatt 1159
 Cp 2126 cttaataatattgaaagaaaggaagggatagatattcaacatttcggtgcacctatt 2067
 Db 1160 ccttttttgcgagcattttgacctctgtttttgctgacccagaaacgctgtgaaacta 1219
 Cp 2066 ccttttttgcgagcattttgacctctgtttttgctgacccagaaacgctgtgaaacta 2007
 Db 1220 aagatgttgagatcatcattggatgacacaaataggtttacatagaactggtatcaacagc 1279
 Cp 2006 aagatgttgagatcatcattggatgacacaaataggtttacatagaactggtatcaacagc 1947
 Db 1290 gtttaagatcctttgagaggttttcgcccgaagaagaggttttccaatgagacattttaaa 1339
 Cp 1946 gtttaagatcctttgagaggttttcgcccgaagaagaggttttccaatgagacattttaaa 1887
 Db 1340 atttcgctatgtgagcgggtattatccctgattgacccgggcaagagaactcggtcgc 1399
 Cp 1886 atttcgctatgtgagcgggtattatccctgattgacccgggcaagagaactcggtcgc 1827
 Db 1400 gcatatactatttcgaataaacttgaattgaattgaattgaattgaattgaattgaatt 1459
 Cp 1824 gcatatactatttcgaataaacttgaattgaattgaattgaattgaattgaattgaatt 1767
 Db 1460 aagatgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1519
 Cp 1766 aagatgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1707
 Db 1520 gtttgcgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1579
 Cp 1706 gtttgcgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1647
 Db 1580 aacatgagagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1639
 Cp 1646 aacatgagagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1587
 Db 1640 ccaaacgacgagcgtgac 1699
 Cp 1546 ccaaacgacgagcgtgac 1527
 Db 1700 ttaactgagagac 1759
 Cp 1526 ttaactgagagac 1467
 Db 1760 gtttgcgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1819
 Cp 1466 gtttgcgagatgagagatgagagatgagagatgagagatgagagatgagagatgagagat 1407
 Db 1820 aacttcgagac 1879
 Cp 1406 aacttcgagac 1347
 Db 1880 aagccctcccgatcatttatttatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1939
 Cp 1346 aagccctcccgatcatttatttatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1287
 Db 1940 aatagacagatcatttatttatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1999
 Cp 1286 aatagacagatcatttatttatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1227
 Db 2000 gtttactcatatatacttatttatttatttatttatttatttatttatttatttatttatttattt 2059

Cp 1226 gtttactcaatatacttatttatttatttatttatttatttatttatttatttatttatttattt 1167
 Db 2060 gtaagatcctttttgataatctcatgacccaaatcccttaacatgaggttttttctctccac 2119
 Cp 1166 gtaagatcctttttgataatctcatgacccaaatcccttaacatgaggttttttctctccac 1107
 Db 2120 tgagcgcagaccccttgagaaagatcaaaagatctctcttaagatcttttttctctccac 2179
 Cp 1106 tgagcgcagaccccttgagaaagatcaaaagatctctcttaagatcttttttctctccac 1047
 Db 2180 gtaatctgctgtctgcaaaacaaaacaccccgctacacacacacacacacacacacacacacac 2239
 Cp 1046 gtaatctgctgtctgcaaaacaaaacaccccgctacacacacacacacacacacacacacacac 987
 Db 2240 caagagctaccacactcttttccgaagtaactgagcttcaacagacacacacacacacacacacac 2299
 Cp 986 caagagctaccacactcttttccgaagtaactgagcttcaacagacacacacacacacacacacac 927
 Db 2300 actgctcttctagctagcgttagtttagccacacacacacacacacacacacacacacacacacacac 2359
 Cp 926 actgctcttctagctagcgttagtttagccacacacacacacacacacacacacacacacacacacac 867
 Db 2360 acatacctcctcctgcttaactctgtttacacagcagctgagctgagctgagctgagctgagctgagct 2419
 Cp 866 acatacctcctcctgcttaactctgtttacacagcagctgagctgagctgagctgagctgagctgagct 807
 Db 2420 cttaacgggttgagctcaagacgataatttaccgagataaagcagcagcagcagcagcagcagcagcagc 2479
 Cp 806 cttaacgggttgagctcaagacgataatttaccgagataaagcagcagcagcagcagcagcagcagcagc 747
 Db 2480 ggggttctgctgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2539
 Cp 746 ggggttctgctgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 687
 Db 2540 cagcgtgagcattgagaaac 2599
 Cp 686 cagcgtgagcattgagaaac 627
 Db 2600 gtaagcggcagc 2659
 Cp 626 gtaagcggcagc 567
 Db 2660 tatcttatacctcctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2719
 Cp 566 tatcttatacctcctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 507
 Db 2720 tgcgcagcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2779
 Cp 506 tgcgcagcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 447
 Db 2780 gcttttctgctgagc 2839
 Cp 446 gcttttctgctgagc 387
 Db 2840 aacctattaccccttctgagc 2899
 Cp 386 aacctattaccccttctgagc 327
 Db 2900 agcgagtcagtcagc 2927
 Cp 326 agcgagtcagtcagc 289

RESULT 7

ID T29158 standard: DNA: 2927 BP.

AC T29158:

DT 09-JAN-1997 (first entry)

DE Plasmid pTRP.

KW Polymerase chain reaction, PCR, amplify; human alanine aminotransferase;

KW ALT: N-terminal region; liver disease; ss.

OS Synthetic.

FH Key Location/Qualifiers

[illegible]

Db 2416 caaacaagcttgcgaacattataactggcggaactactactactagcttccgcgcaacaat 2475
 Cp 1546 CAACAACGCTTGGCAACATTAACTGGCAACTACTACTACTAGCTTCCGCGCAACAT 1487
 Db 2476 caataagatgagatggaggggataaaattgagagacacactctcgcgtcgccctccgg 2535
 Cp 1486 TAATAGATGATGAGAGGGGATAAATATGAGAGGATGATGAGAGGATGATGAGAGG 1427
 Db 2536 ctgctgcttcttctgctgaataaactgagagcggtgagcgatggtctcgcgggtatcattg 2595
 Cp 1426 CTGCGTCTGTTTATTTCTGATTAATCTGAGCGGCTGAGCGGTCTCGCGGTATCATTTG 1367
 Db 2596 cagcactggggcccaagatgtagcgcctccctctatctgtagttatctacacgacggggagtc 2655
 Cp 1366 CAGCACTTGGGACATGTAAGAGGCTTCTGATGAGTATTAACACGACGCGGAGTC 1307
 Db 2655 aggaactatgatagaacaaatagacagatcgtgtgagtaggtgctcactgattaaagc 2715
 Cp 1306 AGGCAACTATGATGAGGAGGAGTAAATAGACAGATCGGTGAGATAGTGCCTCACTGATTAGC 1247
 Db 2716 attgtaactctcagacaaacttactcattatatactatatactattagattgattaaactcatt 2775
 Cp 1246 ATTGTACATGACAGCAAGTTTACTCTATATATACATTAGATTGATTTAAACATTCATT 1187
 Db 2776 tttaattttaaaggatctaggtgagatccctcttctgataactctcattacacaaatccctt 2835
 Cp 1186 TTTAATTTAAATGAACTAGTAAATGATCTTTTTCATATATCTATGACCAAAATCCCTT 1127
 Db 2836 agcgtgaattctgtctcactgagcgctcagaccccgtagaagaagatcaaaaggatctctt 2895
 Cp 1126 AAGCGATTTTCTTTCTACTGAGCGCTCAGACGCGGTAGAAAGATCAAGAGATCTCTT 1067
 Db 2895 gagatccctctttctgctgaatctgctgcttgcgaacaaacacacccacccgctacag 2955
 Cp 1066 CAGATTTTCTTTCTACTGAGCGCTCAGACGCGGTAGAAAGATCAAGAGATCTCTT 1007
 Db 2955 aggtgatttctgctgagatcaaaagctacacaaactcttcttccaaaggaactggtctca 3015
 Cp 1007 CAGATTTTCTTTCTACTGAGCGCTCAGACGCGGTAGAAAGATCAAGAGATCTCTT 947
 Db 3016 cagagcgagatcacaact 3075
 Cp 946 CAG 887
 Db 3075 agaatctgtgagacgcctcactacacctcgcctgctgctgctgctgctgctgctgctg 3135
 Cp 886 AGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 827
 Db 3135 ccagtgagagataagtcgtctctctctctctctctctctctctctctctctctctct 3195
 Cp 826 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 767
 Db 3195 ccaacgctcggctggaacggggtctctctcacaacccagcttggagcgacacact 3255
 Cp 766 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707
 Db 3255 acacgaactatgatacctacaggtgagctatagaaagggcgacgcttcccaagagga 3315
 Cp 706 ACACGAACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647
 Db 3315 aaaggggagagaggttatccggttaagcttctctctctctctctctctctctctctct 3375
 Cp 646 TAAAG 587
 Db 3375 tccaggggggaacacccctgggtatcttattagctcctgctgctgcttctcgcacactctgacttg 3435
 Cp 587 TCTAG 527
 Db 3435 agcctcgtatttttctgagctcgtcagggggcgagagcctatgaaaaaccccgagcag 3495
 Cp 526 AAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 467

Db 3496 cggccttttaccggttcctggccttttgcgacgttttgcgacgttttgcgacgttttgcgacgt 3555
 Cp 466 CGGCTTTTACGGTTCCTGGGCTTTTGTGAGCTTTTGTGAGCTTTTGTGAGCTTTTGTGAGCT 407
 Db 3556 tatccctgattctcttggtataacgcgtattacgcgttttgcgacgttttgcgacgttttgcgacgt 3615
 Cp 406 TATCCCTGATTCTGTGATACCGCTATTACCGCTTTTGTGAGCTTTTGTGAGCTTTTGTGAGCT 347
 Db 3615 gcagccgagacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 3664
 Cp 346 GCAGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 298

RESULT 11

ID Q79545 standard; DNA; 6714 BP.
 AC Q79545.
 DT 10-AUG-1995 (first entry)
 DE pWB254b plasmid expression vector encoding Kientaq-278.
 KW Thermostable polymerase; Kientaq-278; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT cds 1..1665
 FT /*tag= a
 FN WC9426766-A.
 PD 24-NOV-1994.
 PF 22-FEB-1994; U01867.
 PR 19-FEB-1993; US-021623.
 FR 22-FEB-1994; US-021623.
 PA (BARN/) BARNES W M.
 PI Barnes WM.
 DR WPI: 95-006692/01.
 DR P-PSDB: R66209.
 PT DNA polymerase and formulations comprising it - allowing the
 PT amplification of sequences up to 35 kilobases and reducing the
 PT mutagenicity generated by the PCR process.
 PS Disclosure: Page 39-44, 79pp. English.
 CC Kientaq-278 is a thermostable polymerase lacking 3'-exonuclease
 CC activity. The AA sequence is substantially the same AA sequence as
 CC Thermus aquaticus DNA polymerase but lacks the N-terminal 280 AAs. A
 CC DNA encoding such a polymerase is claimed, as is the polymerase
 CC itself. The polymerase also has substantially the same AA sequence
 CC as that of Thermus flavus, but excludes the N terminal 279 AAs.
 CC Primers Q79543, Q79544, Q79553 and Q79554 can be used for the
 CC amplification of the gene for Kientaq-278. Essentially the same
 CC primers can be used for the amplification of the analogous gene from
 CC Thermus flavus (see Q79543, Q79556, Q79557). An initiator Met and a
 CC Gly occupy the first two N-terminal posns of Kientaq-278.
 CC Previously occupied by residues 279 and 280 of T. aquaticus
 CC DNA polymerase, followed by the AA sequence of wt T. aquaticus
 CC DNA polymerase, beginning with the AA residue at posn. 281.
 CC Plasmid pWB254b contains the sequence in Q79545.
 SS Sequence 6714 BP. 1489 A. 1862 C. 1842 G. 1521 T.

Query Match 70.6%; Score 2375; DB 13; Length 6714;
 Best Local Similarity 99.5%; Pred. No. 0.00e+00;
 Matches 2393; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

Db 1819 atagtagcgccctgtacgagcattaaagcgaggggtgtaggttgcgacgagct 1878
 Cp 2690 ATGGGAGCGGCGCTGTAGCGGCGGTATTAAGCGCGCGCGGTGTAGCGCGCGGTGTAGCGCGCGGT 1631
 Db 1879 gacggtatcacttgccagagcgccttagcgccgctctctctctctctctctctctctctctctctctctct 1938
 Cp 2630 GATCGGTACACTGCTATGCGGTATAGCGCGCGGTGTAGCGCGCGGTGTAGCGCGCGGTGTAGCGCGCGGT 1571
 Db 1939 cgcacgttcgcgggcttcccccgcacaaactctaaacagagctccctctttaggtctcag 1998
 Cp 2570 CGCCAGCTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGT 2511
 Db 1999 attagtgctttacgacgacccctgaccccccaaaacttgaattggatgaggttcacgtag 2058
 Cp 2510 ATTAACTGCTTACGCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGT 2451

DB 2059 taagcgcacccctcaatcaaaagcttttttccgcttttttcaatttttcaatttttcaattttttttaa 2118
CP 2450 tggagcattggcctgctgataacgagtttttttggccttttttgaatttttttgaatttttttttaa 2361
DB 2139 taatgaatttttttttccaaaatttaaaacacatcaaaatttttttttttttttttttttttttt 2178
CP 2390 taatggcatttttttttcaaaatt 2331
DB 2179 tttataaaggaatt 2398
CP 2330 tttataaaggaatt 2271
DB 2239 atttaaaaggaatt 2298
CP 2270 atttaaaaggaatt 2212
DB 2299 taatgaatt 2358
CP 2211 taatgaatt 2152
DB 2359 caatgaacgaatt 2418
CP 2151 taatgaacgaatt 2092
DB 2419 taatgaatt 2478
CP 2091 taatgaatt 2032
DB 2479 taatgaacgaatt 2538
CP 2031 taatgaacgaatt 1972
DB 2539 taatgaacgaatt 2598
CP 1971 taatgaacgaatt 1912
DB 2599 ttttttaatt 2658
CP 1911 ttttttaatt 1852
DB 2659 taatgaacgaatt 2718
CP 1851 taatgaacgaatt 1792
DB 2719 taatgaacgaatt 2778
CP 1791 taatgaacgaatt 1732
DB 2779 taatgaacgaatt 2838
CP 1731 taatgaacgaatt 1672
DB 2839 taatgaacgaatt 2898
CP 1671 taatgaacgaatt 1612
DB 2899 taatgaacgaatt 2958
CP 1611 taatgaacgaatt 1552
DB 2959 taatgaacgaatt 3018
CP 1551 taatgaacgaatt 1492
DB 3019 taatgaacgaatt 3078
CP 1491 taatgaacgaatt 1432
DB 3079 taatgaacgaatt 3138
CP 1431 taatgaacgaatt 1372
DB 3139 taatgaacgaatt 3198

CP 1371 catttgacgacacttgaggaac 1432
DB 3199 gaatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3258
CP 1311 gactcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 1252
DB 3259 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3318
CP 1251 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 1192
DB 3319 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3378
CP 1191 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 1132
DB 3379 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3438
CP 1131 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 1072
DB 3439 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3498
CP 1071 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 1012
DB 3499 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3558
CP 1011 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 952
DB 3559 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3618
CP 951 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 892
DB 3619 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3678
CP 891 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 832
DB 3679 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3738
CP 831 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 772
DB 3739 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3798
CP 771 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 712
DB 3799 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3858
CP 711 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 652
DB 3859 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3918
CP 651 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 592
DB 3919 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 3978
CP 591 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 532
DB 3979 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4038
CP 531 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 472
DB 4039 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4098
CP 471 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4158
DB 4099 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4198
CP 411 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 352
DB 4159 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4218
CP 351 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 292
DB 4219 taatcagcgaactatgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaatgaacgaat 4224
CP 292

Cp 291 AAIACG 286

RESULT 12

ID Q62676 standard: J. DNA. 3832 BP.

AC Q62676

DI 11-JAN-1995 (first entry)

DE Plasmid pASK60-Strep encodes streptavidin-binding peptide.

KW Streptavidin binding peptide; fusion protein; pASK60-Strep;

KW affinity chromatography; purification; peptide tag; detection: ds.

OS Synthetic.

FH Key Location/Qualifiers

FT cds 1..3831

FT /tag= a

FT /label= Reading_Frame_a

FT /note= "contains many termination codons"

FT /tag= b

FT /label= Reading_Frame_b

FT /note= "contains many termination codons"

FT /tag= c

FT /label= Reading_Frame_c

FT /note= "contains many termination codons"

FT /tag= d

FT /label= lacZ_mini_cistron

FT /note= "reading frame a"

FT /tag= e

FT /label= lacZ_mini_cistron

FT /note= "reading frame a"

FT /tag= f

FT /label= lacZ_mini_cistron

FT /note= "reading frame a"

FT /tag= g

FT /label= OmpA_signal_peptide

FT /note= "reading frame b"

FT /tag= h

FT /label= OmpA_signal_peptide

FT /note= "reading frame b"

FT /tag= i

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= j

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= k

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= l

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= m

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= n

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= o

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= p

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= q

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= r

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= s

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= t

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= u

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= v

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= w

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= x

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

FT /tag= y

FT /label= OmpA_signal_peptide

FT /note= "reading frame c"

工
L
Z
U
Q
Y

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

>US-09-020-716-1
Description:
(1-3363) from US09020716 seq
3363
Perfect Score:
N.A. Sequence:
Comp:
1 TCGACCTCGAGGNGGAGGCGG...
AGTGGAGGCTGCTGGGCTGGG...
GGGAGGAGCAGCCAGCTATCC
GGCTTTGTGTGCTGGCGATGG
3363

```

Scoring table. TABLE default.

Nmatch STD : Dbase 0: Query 0

Searched

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: embl-est56

```
Database:
genbank-est|09
sssi_wm:4:am_qss_
sssi_wm:5:am_qss_
sssi_wm:7:am_qss_
sssi_wm:11:am_qss_
```

10:gb-est16 7:gb-est11 8:gb-est12 9:gb-est13
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17
14:gb-est18 15:gb-est19 16:gb-est2 17:gb-est20
18:gb-est21 19:gb-est22 20:gb-est23 21:gb-est24
22:gb-est25 23:gb-est26 24:gb-est27 25:gb-est28
26:gb-est29 27:gb-est30 28:gb-est31 29:gb-est32
30:gb-est33 31:gb-est34

Statistics: Mean 12.051, Variance 1.914, scale 6.295

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			Length	DR	ID	Description	Pred. No.
	Score	Match						
1	54	19.7	711	26	AS004508	Homo sapiens	genomic D	0.00e+00
2	54	19.7	711	27	AG010947	Homo sapiens	genomic D	0.00e+00
3	57	19.5	721	27	AG011001	Homo sapiens	genomic D	0.00e+00
4	48	19.3	721	26	AC004562	Homo sapiens	genomic D	0.00e+00
5	65	18.9	718	26	AG004363	Homo sapiens	genomic D	0.00e+00
6	65	18.9	718	27	AG010489	Homo sapiens	genomic D	0.00e+00
7	53	18.7	648	26	AG003787	Homo sapiens	genomic D	0.00e+00
8	50	18.7	698	27	AG003976	Homo sapiens	genomic D	0.00e+00
9	64	18.6	696	27	AG003765	Homo sapiens	genomic D	0.00e+00
10	54	18.6	696	26	AG003576	Homo sapiens	genomic D	0.00e+00
11	54	18.6	720	28	AG013858	Homo sapiens	genomic D	0.00e+00
12	54	18.6	720	26	AG004562	Homo sapiens	genomic D	0.00e+00
13	51	18.4	747	26	AG007052	Homo sapiens	genomic D	0.00e+00

C	14	609	18.1	644	28	A015013	Git-HSP-2374C4.1F	Git	0.00e+00
	15	603	17.9	692	27	A0030357	Git-HSP-2374K1.1F	Git	0.00e+00
C	16	603	17.9	692	27	A0030357	Homo sapiens genomic D	0.00e+00	
	17	603	17.8	703	26	A0011761	Homo sapiens genomic D	0.00e+00	
C	18	598	17.8	748	17	A0114281	1.59 Dirosophila p-12	D	0.00e+00
	19	595	17.7	654	26	B78704	Citrus-SJA-A-8905.1PC	C	0.00e+00
C	20	594	17.7	654	27	A5094454	Homo sapiens genomic D	0.00e+00	
	21	593	17.6	640	27	A0014338	Git-HSP-2374E1.1F	Git	0.00e+00
C	22	593	17.5	652	28	A0118633	Git-HSP-2374M1.1F	Git	0.00e+00
	23	588	17.3	594	26	C1786	Clona intestinalis gen	0.00e+00	
C	24	583	17.2	666	27	A0030909	Git-HSP-2374E4.1F	Git	0.00e+00
	25	577	17.2	689	28	A00443394	Homo sapiens genomic D	0.00e+00	
C	26	566	16.9	642	29	A0113213	Git-HSP-2374K2.1F	Git	0.00e+00
	27	566	16.8	682	27	A0014543	Git-HSP-2374K3.1F	Git	0.00e+00
C	28	557	16.6	615	28	A0113424	Git-HSP-2374E2.1F	Git	0.00e+00
	29	559	16.6	742	28	A0116342	Git-HSP-2374E3.1F	Git	0.00e+00
C	30	551	16.4	604	25	A0047705	Homo sapiens genomic D	0.00e+00	
	31	550	16.4	634	27	A0048010	Git-HSP-2374F.1F	Git	0.00e+00
C	32	549	16.3	657	27	A0017804	Git-HSP-2374K1.1F	Git	0.00e+00
	33	544	16.2	633	27	A0040787	Git-HSP-2374K2.1F	Git	0.00e+00
C	34	545	16.2	661	27	A0057503	Git-HSP-2374M2.1F	Git	0.00e+00
	35	545	16.2	690	27	A00393399	RPC111-39F.1F	RPC111	0.00e+00
C	36	540	16.1	637	27	A0037360	Git-HSP-2365E1.1F	Git	0.00e+00
	37	534	15.9	594	26	B71365	RPC111-9N1	TP	0.00e+00
C	38	534	15.9	626	28	A0111347	Git-HSP-2374M1.1F	Git	0.00e+00
	39	531	15.8	549	27	A0006191	RPC111-10E5	TPB	0.00e+00
C	40	530	15.8	630	27	A0030239	RPC111-2313	TPB	0.00e+00
	41	538	15.8	743	25	A0014824	Homo sapiens genomic D	0.00e+00	
C	42	537	15.7	624	27	A0041367	Homo sapiens genomic D	0.00e+00	
	43	535	15.6	540	27	A00474307	Git-HSP-2382N.1F	Git	0.00e+00
C	44	536	15.6	594	25	A0047826	Homo sapiens genomic D	0.00e+00	
	45	535	15.6	600	28	A0111411	Git-HSP-2373F3.1F	Git	0.00e+00

ALIGNMENTS

RESULT	LOCUS	DEFINITION	711 bp	DNA	GSS	29-JAN-1998
1	AG004508	Homo sapiens genomic DNA, 21q region, survey sequence.			S594BG38, genomic	
	AG004508	Accession				
	NID	q2822058				
	KEYWORDS	GSS				
	SOURCE	Homo sapiens DNA, clone:S594BG38.				
	ORGANISM	Homo sapiens				
	REFERENCE	Eukaryotes: Metazoa, Chordata; Vertebrata: Mammalia: Eutheria: Primates: Catarrhini, Hominoidea, Homo.				
	AUTHORS	1 (bases 1 to 711)				
	TITLE	Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
	JOURNAL	Homo sapiens genomic DNA, chromosome 21q				
	REFERENCE	Published Only in Database (1998) In press				
	AUTHORS	2 (bases 1 to 711)				
	TITLE	Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
	JOURNAL	Direct Submission				
	REFERENCE	Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.				
	AUTHORS	Masahiro Hattori, Kitasato University, Department of Science, Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan				
	TITLE	(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, tel:0427-78-5732, Fax:0427-78-9561)				

FEATURES

source

00000

COUNT

Z

ry. Matu-

t. Loca.

2105

1888

US-09-020-716-1.1st

Wed Mar 17 09:43:26 1999

```

REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-Jul-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES             Location/Qualifiers
     source           1..721
     map              195 a 163 c 175 g 179 t 9 others
     base count       195 a 163 c 175 g 179 t 9 others
     origin            Query Match 19.5%; Score 657; DB 27; Length 721;
                    Best Local Similarity 98.7%; Pred. No. 0.00e+00;
                    Matches 583; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

Db 32 CATCTTACGGATGGCATGACATAGAGAATTTATGCAAGTCTGCTGCCATACCATGAGTGT 91
Cp 1772 CATCTTACGGATGGCATGACATAGAGAATTTATGCAAGTCTGCTGCCATACCATGAGTGT 1713
Db 92 AACACTCGCGCCAACTTACTTCTGACAGCATGCGAGGACCGAGGAGGAGTAAACCGCTTTT 151
Cp 1712 AACACTCGCGCCAACTTACTTCTGACAGCATGCGAGGACCGAGGAGGAGTAAACCGCTTTT 1653
Db 152 TTGCACAAACAGCAGGAGTGCACAGCATGCGCTGTAGCAATGGCAACAACTTGGCG 271
Cp 1592 TTGCACAAACAGCAGGAGTGCACAGCATGCGCTGTAGCAATGGCAACAACTTGGCG 1593
Db 272 AACACTTAACTTGGCGAACTACTTCTAGCTTCCGCGCAACATTAATAGCTGGATG 331
Cp 1532 AACACTTAACTTGGCGAACTACTTCTAGCTTCCGCGCAACATTAATAGCTGGATG 1473
Db 332 GAGGCGGATAAAGTTGCGAGGAGTCTGCGCTCGGCTCGGCTCGGCTGGTTTATT 391
Cp 1472 GAGGCGGATAAAGTTGCGAGGAGTCTGCGCTCGGCTCGGCTCGGCTGGTTTATT 1413
Db 392 GCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGGCTATCATTCAGTGGGCGCA 451
Cp 1412 GCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGGCTATCATTCAGTGGGCGCA 1353
Db 452 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGGAGTACGGCAACTATGGAT 511
Cp 1352 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGGAGTACGGCAACTATGGAT 1293
Db 512 GAACGAAATAGACAGATCGGTGACATAGCTGCTCACTGATTAAAGC-TTGGTAAGTCA 570
Cp 1292 GAACGAAATAGACAGATCGGTGACATAGCTGCTCACTGATTAAAGC-TTGGTAAGTCA 1233
Db 571 GACCAAGTTTACTG-TATATACCTTTAGATTTTAAACTTCATTTTAAATTTAAAGG 629
Cp 1222 GACCAAGTTTACTGATATATATCTTTAGATTTTAAACTTCATTTTAAATTTAAAGG 1173
Db 630 ATCTAGGTGAAGATCCTTTT-GATAATCNCATGACCAAA-TCCCTTAAGCTGAGTTTCG 687
Cp 1172 ATCTAGGTGAAGATCCTTTTTCATATCTCAAGCAAAATCCCTTAAGCTGAGTTTCG 1113
Db 588 TTCCACTGAGCTGCAGACCCGCTAGAAAAAT 719
Cp 1112 TTCCACTGAGCTGCAGACCCGCTAGAAAAAT 1081

```

```

RESULT 4
LOCUS AG04662 721 bp DNA GSS 30-JAN-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic survey sequence.
ACCESSION AG04662
NID Q2826191
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 879G12X91.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES             Location/Qualifiers
     source           1..721
     map              195 a 163 c 175 g 179 t 9 others
     base count       195 a 163 c 175 g 179 t 9 others
     origin            Query Match 19.5%; Score 657; DB 26; Length 721;
                    Best Local Similarity 98.7%; Pred. No. 0.00e+00;
                    Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

Db 32 CATCTTACGGATGGCATGACATAGAGAATTTATGCAAGTCTGCTGCCATACCATGAGTGT 91
Cp 1772 CATCTTACGGATGGCATGACATAGAGAATTTATGCAAGTCTGCTGCCATACCATGAGTGT 1713
Db 92 AACACTCGCGCCAACTTACTTCTGACAGCATGCGAGGACCGAGGAGGAGTAAACCGCTTTT 151
Cp 1712 AACACTCGCGCCAACTTACTTCTGACAGCATGCGAGGACCGAGGAGGAGTAAACCGCTTTT 1653
Db 152 TTGCACAAACAGCAGGAGTGCACAGCATGCGCTGTAGCAATGGCAACAACTTGGCG 271
Cp 1592 TTGCACAAACAGCAGGAGTGCACAGCATGCGCTGTAGCAATGGCAACAACTTGGCG 1593
Db 272 AACACTTAACTTGGCGAACTACTTCTAGCTTCCGCGCAACATTAATAGCTGGATG 331
Cp 1532 AACACTTAACTTGGCGAACTACTTCTAGCTTCCGCGCAACATTAATAGCTGGATG 1473
Db 332 GAGGCGGATAAAGTTGCGAGGAGTCTGCGCTCGGCTCGGCTCGGCTGGTTTATT 391
Cp 1472 GAGGCGGATAAAGTTGCGAGGAGTCTGCGCTCGGCTCGGCTCGGCTGGTTTATT 1413
Db 392 GCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGGCTATCATTCAGTGGGCGCA 451
Cp 1412 GCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGGCTATCATTCAGTGGGCGCA 1353
Db 452 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGGAGTACGGCAACTATGGAT 511
Cp 1352 GATGGTAAGCCCTCCCGTATCGTAGTTATCTACAGCGGGGAGTACGGCAACTATGGAT 1293
Db 512 GAACGAAATAGACAGATCGGTGACATAGCTGCTCACTGATTAAAGC-TTGGTAAGTCA 570

```



```

DB 604 TAAAGTTTAAATTT 614
      |||||||
CP 2245 TAAAGTTTAAATTT 2246

RESULT 15
AC004352 692 bp DNA 5SS 10-III-1998
DEFINITION chr.usp.2327k21 repeat map from human genome clone 2327k21,
genomic survey sequence.
AC004352
MIM 43406184
KEYWORDS 5SS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHOR Adams M.D., Boumsley S.D., Field C.F., Bass S., Linber K.,
Golden K., Berry K., Granger D., Suk E., Wible C., Shizuya H.,
Simon M. and Venter J.C.
TITLE Use of a random BAC End Sequences Database for Sequence-Ready Map
Building
JOURNAL Published (1997)
JOURNAL PUBMED 9094353
OTHER_IDS: C11-HSP-2327K21.TR
CONTACT: Mark Adams
DEPARTMENT of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
TEL: 301 838 0200
FAX: 301 838 0208
EMAIL: mdadams@tigr.org
NOTES: Genomes are available from Research Genetics (info@resgen.com) BAC
end search page:
http://www.tigr.org/Research/BAC_end_search.html.
Seq primer: M13-21
CLASS: BAC ends.
FEATURES
SOURCE
1..692
location/qualifiers
/orqanism="Homo sapiens"
/notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII"
zdb_xref="taxon:9606"
zdb_xref="2327K21"
/clone_lib="C11-HSP"
/sex="Male"
/cell_type="Sperm"
BASE COUNT 158 a 193 c 186 g 155 t
CPISIN

Query Match 17.9% Score 604; DB 27; Length 692;
Res. Local Similarity 99.84; Pred. No. 0.09e-03;
Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 88 GCTGGGATATCATGCTATATGCTGTTCTGCTGTAATTTATCGGTCACAAATTC 147
      |||||||
CP 2245 GCTGGGATATCATGCTATATGCTGTTCTGCTGTAATTTATCGGTCACAAATTC 2246

DB 148 TACACAAATATGAGAGCGGAATATAAGCTTAAAGCTTGGCTGCTTAATGATGAGCT 207
      |||||||
CP 2245 TACACAAATATGAGAGCGGAATATAAGCTTAAAGCTTGGCTGCTTAATGATGAGCT 207

DB 122 CATACAAATATGAGAGCGGAATATAAGCTTAAAGCTTGGCTGCTTAATGATGAGCT 181
      |||||||
CP 2245 CATACAAATATGAGAGCGGAATATAAGCTTAAAGCTTGGCTGCTTAATGATGAGCT 181

DB 208 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
      |||||||
CP 2245 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267

DB 142 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
      |||||||
CP 2245 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241

DB 268 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
      |||||||
CP 2245 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327

DB 242 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
      |||||||
CP 2245 AATTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

DB 428 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
      |||||||
CP 2245 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387

DB 402 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
      |||||||
CP 2245 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361

```

```

DB 388 CTCATCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 447
      |||||||
CP 2245 CTCATCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 447

DB 448 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 507
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 507

DB 508 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 567
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 567

DB 568 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 627
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 627

DB 628 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 687
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 687

DB 688 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 747
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 747

DB 748 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 807
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 807

DB 808 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 867
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 867

DB 868 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 927
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 927

DB 928 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 987
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 987

DB 988 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1047
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1047

DB 1048 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1107
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1107

DB 1108 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1167
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1167

DB 1168 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1227
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1227

DB 1228 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1287
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1287

DB 1288 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1347
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1347

DB 1348 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1407
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1407

DB 1408 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1467
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1467

DB 1468 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1527
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1527

DB 1528 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1587
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1587

DB 1588 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1647
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1647

DB 1648 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1707
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1707

DB 1708 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1767
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1767

DB 1768 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1827
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1827

DB 1828 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1887
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1887

DB 1888 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1947
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1947

DB 1948 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2007
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2007

DB 2008 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2067
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2067

DB 2068 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2127
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2127

DB 2128 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2187
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2187

DB 2188 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2247
      |||||||
CP 2245 TGTGAGCAAAAGGCGGTAATAGGTAATCAATCAATCAATCAATCAATCAATCAATCAAT 2247

```

Search completed: Sun Mar 14 09:38:11 1999
Job time : 5791 secs.

QY 1755 CATGCCATCGTAAGATCTTTTGTGACATGATGATCAACCAAGTCATCTCAGA 1814
 DB 2103 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2162
 QY 1815 AAGAGATGATGAGGATGATCTTCTGCGGCGGCGGCAATATACCGCGCC 1874
 DB 2163 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2222
 QY 1875 ACATAGCAGAACTTTAAAGTGTCTATCTTGGAAAGCTTCTGCGGCGGCGGCAATATACCGCGCC 1934
 DB 2223 AAGATGATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2282
 QY 1945 AAGATGATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 1994
 DB 2283 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2342
 QY 1995 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2054
 DB 2343 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2402
 QY 2055 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2114
 DB 2403 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2462
 QY 2115 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2174
 DB 2463 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2522
 QY 2175 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2234
 DB 2523 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2582
 QY 2235 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2294
 DB 2583 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2642
 QY 2295 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2354
 DB 2643 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2702
 QY 2355 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2414
 DB 2703 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2762
 QY 2415 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2474
 DB 2763 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2822
 QY 2475 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2534
 DB 2823 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2882
 QY 2535 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2594
 DB 2883 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2942
 QY 2595 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2654
 DB 2943 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 3002
 QY 2655 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2714
 DB 3003 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 3062
 QY 2715 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2774
 DB 3063 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 3122
 QY 2775 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 2834
 DB 3123 ATATGTATGCGGACGACATCTTCTGCGGCGGCGGCAATATACCGCGCC 3182

QY 2835 ACGGCGATGAGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 2894
 DB 3183 GGC 3185
 QY 2895 GGC 2897
 RESULT 4
 LOCUS U02449 2967 bp DNA Circular SYN 29-MAR-1996
 DEFINITION Cloning vector pDIRECT, complete sequence.
 ACCESSION U02449
 NID 9413815
 KEYWORDS
 SOURCE Cloning vector pDIRECT.
 ORGANISM Cloning vector pDIRECT.
 REFERENCE 1 (bases 1 to 2967)
 AUTHORS Kitts, P.A.
 TITLE CLONTECH Vectors On Disc version 1.3
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2967)
 AUTHORS Kitts, P.A.
 TITLE Direct Submission
 JOURNAL Submitted (09-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
 COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3.
 This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH. This vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail: TECH@CLONTECH.COM.
 FEATURES
 Location/Qualifiers
 1..2967
 /organism="Cloning vector pDIRECT"
 /base_xref="taxon:31799"
 BASE COUNT 726 a 756 c 735 g 750 t
 ORIGIN
 Query Match. 85.5%, Score 2878, EB 32, Length 2967.
 Best Local Similarity 99.7%; Pred. No 0.00e+00;
 Matches 2901, Conservative 0; Mismatches 5; Indels 3; Gaps 1;
 DB 21 TOTALAGCGCGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 2894
 QY 2835 ACGGCGATGAGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 2894
 DB 81 T---TCATCGCGCGCTGTTTACAGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 137
 CP 2849 GCGCTACGCGCGCTGTTTACAGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 2790
 DB 138 TAATCGCTTTCGACAGCATATCGCGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 197
 CP 2789 TAATCGCTTTCGACAGCATATCGCGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 2730
 EL 196 CGATCGCGCTTTCGACAGCATATCGCGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 257
 QY 2729 CGATCGCGCTTTCGACAGCATATCGCGCTGCTGACCTGCGGAAACCGCTGAGCTACCTAACT 2670
 DB 258 CGCATTAAGCGCGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 317
 CP 2669 CGCATTAAGCGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 2610
 DB 318 CGCATTAAGCGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 377
 CP 2609 CGCATTAAGCGCGGCGGTAATACGACTACATATAGGCGAAATGAGAGTACGAGCGGCT 2650

SOURCE: human.
ORGANISM: Homo Sapiens
REFERENCE: 1 (bases 1 to 6824)
AUTHORS: Galvez, C., Gallo, E., Rossolini, G.M., Ricci, M.L., and Thaler, M.C.
TITLE: Genetic vector for multiple stable integration of DNA sequences into the genome of the yeasts *Kluyveromyces fragilis* and *Saccharomyces cerevisiae* and plasmids containing the same patent. EP 654745 A 1 21 APR 1993.
JOURNAL: SCLAVO S.p.A.
FEATURES: Location/Qualifiers
source 1..6824
 /organism="Homo Sapiens"
 /db_xref="taxon:9606"
BASE COUNT 1814 a 1521 c 1726 g 1763 t
ORIGIN
Query Match 82.9% Score 2796; DB 22; Length 6824;
Best Local Similarity 99.5%; Pred. No. 0.00000;
Matches 2846; Conservative 9; Mismatches 6; Indels 9; Gaps 3;
Db 4996 GCAATCGGCTATAGTACGATGATTAACAT...TACATGCGCTGCTTTTAAAGGTC 3952
|||||
Cp 2481 GCAATCGGCTATAGTACGATGATTAACAT...TACATGCGCTGCTTTTAAAGGTC 2821
Db 4953 GTTACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4012
|||||
Cp 2823 GTTACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2761
Db 4913 GCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4072
|||||
Cp 2760 GCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2701
Db 4973 TCAATGCTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4132
|||||
Cp 2700 TCAATGCTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2644
Db 4933 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4192
|||||
Cp 2644 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2584
Db 4993 TCGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4252
|||||
Cp 2583 TCGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2524
Db 4953 GTTACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4312
|||||
Cp 2523 GTTACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2464
Db 4913 ATGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4372
|||||
Cp 2463 ATGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2404
Db 4973 GCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4432
|||||
Cp 2403 GCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2344
Db 4933 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4492
|||||
Cp 2343 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2284
Db 4993 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4552
|||||
Cp 2283 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2227
Db 4953 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4612
|||||
Cp 2226 TCACTGGTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2167
Db 4913 AATATGATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4672
|||||

Cp 2166 AATATGATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2107
Db 4673 GAAGTATGATGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4742
|||||
Cp 2106 GAAGTATGATGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 2047
Db 4733 CTTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4794
|||||
Cp 2046 CTTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1987
Db 4793 GCTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4852
|||||
Cp 1986 GCTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1927
Db 4853 TCTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4912
|||||
Cp 1926 TCTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1867
Db 4913 ATTATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 4972
|||||
Cp 1866 ATTATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1807
Db 4973 TCACTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5032
|||||
Cp 1806 TCACTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1747
Db 5033 AGAATATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5092
|||||
Cp 1746 AGAATATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1687
Db 5093 AAGATATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5152
|||||
Cp 1686 AAGATATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1627
Db 5153 TCGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5212
|||||
Cp 1626 TCGCTTTCGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1567
Db 5213 CAGCATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5272
|||||
Cp 1566 CAGCATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1507
Db 5273 TCTAGCTTCCCGGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5332
|||||
Cp 1506 TCTAGCTTCCCGGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1447
Db 5333 TCTGCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5392
|||||
Cp 1446 TCTGCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1387
Db 5393 TGGCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5452
|||||
Cp 1386 TGGCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1327
Db 5453 TATCTACATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5512
|||||
Cp 1326 TATCTACATGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1267
Db 5513 AGTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5572
|||||
Cp 1266 AGTCTGCTGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1207
Db 5573 GATTCATTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5632
|||||
Cp 1206 GATTCATTAAAGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1147
Db 5633 TCTCAAGCAAAATTCCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5692
|||||
Cp 1146 TCTCAAGCAAAATTCCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1087
Db 5693 AAAGATCAAGGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 5752
|||||
Cp 1086 AAAGATCAAGGCTGCTGCTTACCACTTAATGCGCTTGCAGCAGATCGGCTTTCG 1027

[illegible]

Wed Mar 17 09:43:27 1999

REFERENCE 2 (bases 1 to 143)
 AUTHORS Moss, B., Elroy-Stein, O., Mizukami, T., Alexander, W. A. and
 Fuerst, R. K.
 TITLE Product review. New mammalian expression vectors
 JOURNAL Nature 348 (6296), 91-92 (1990)
 MEDLINE 91043055
 REFERENCE 3 (bases 1 to 5845, 8161 to 11271, 5846 to 8160)
 AUTHORS Chen, X., Li, Y., Xiong, K. and Wanner, I. F.
 TITLE A self-initiating eukaryotic transient gene expression system based
 on cotransfection of bacteriophage T7 RNA polymerase and DNA
 vectors containing a T7 autogene
 JOURNAL Unpublished
 REFERENCE 4 (bases 1 to 11271)
 AUTHORS Chen, X.
 TITLE Direct Submission
 JOURNAL Submitted (15-APR-1994) Xiaozhuo Chen, Molecular Biology, Edison
 Biotechnology Institute, 102 Wilson Hall, West Green, Ohio
 University, Athens, Ohio, 45701, USA
 FEATURES
 source
 L..11271
 /organism="Escherichia coli"
 /strain="HM174"
 /db_xref="taxon:562"
 /clone="pT77/I7h5H"
 BASE COUNT 2856 a 2777 c 2799 g 2839 t
 ORIGIN
 Query Match 81.3% Score 2737; DB 16, Length 11271;
 Best Local Similarity 99.7%; Pred. No. 0.00e+00;
 Matches 2774; Conservative 0; Mismatches 1; Indels 6; Gaps 2;
 Db 8491 GGCGTCTGTTTACAAAGTCGTGACTGGGAAACCCCTGGCGTTACCCAACTTATCGCCT 8550
 Cp 2840 GGCGTCTGTTTACAAAGTCGTGACTGGGAAACCCCTGGCGTTACCCAACTTATCGCCT 2781
 Db 8551 TCGAGACATCCCGCTTTCGACGTCGTGATAGCGGTAATAGCGAAGAGCGCCCGATCGGCC 8610
 Cp 2789 TCGAGACATCCCGCTTTCGACGTCGTGATAGCGGTAATAGCGAAGAGCGCCCGATCGGCC 2721
 Db 8611 TCGAGACATCCCGCTTTCGACGTCGTGATAGCGGTAATAGCGAAGAGCGCCCGATCGGCC 8670
 Cp 2720 TCGAGACATCCCGCTTTCGACGTCGTGATAGCGGTAATAGCGAAGAGCGCCCGATCGGCC 2664
 Db 8671 AAGCGCGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8730
 Cp 2663 AAGCGCGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2604
 Db 8731 GCGCGCTCTTTCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 8790
 Cp 2603 GCGCGCTCTTTCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 2544
 Db 8791 AGCTCTAATCGGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8850
 Cp 2543 AGCTCTAATCGGCGGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2484
 Db 8851 GAAAAAATCGATAGGTCGTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8910
 Cp 2483 GAAAAAATCGATAGGTCGTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2424
 Db 8911 TCGCGCTTTCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 8970
 Cp 2423 TCGCGCTTTCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 2364
 Db 8971 AAGACTCAACCTTATCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 9030
 Cp 2363 AAGACTCAACCTTATCGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCT 2304
 Db 9031 CTATTGGTATANAATGACTGATTTTAAACAAAATTTAAACGGAATTTTAAACAAAATTT 9090
 Cp 2303 CTATTGGTATANAATGACTGATTTTAAACAAAATTTAAACGGAATTTTAAACAAAATTT 2244
 Db 9091 AAGGTTTACAAATTTCCCAAGTGGACATTTTCCGCGGAAATTTTCCGCGGAAATTTTCCGCGG 9150

Cp 2243 AACGCTTACAAATTT--AGGTGGCACTTTTCGGGAAATGTGCGCGGACCCCTATTGT 2187
 Db 9151 TTATTTTCTAAATACATTCAAATATATGATCGGCTCATGAGACAATAACCCCTGATAATG 9210
 Cp 2186 TTATTTTCTAAATACATTCAAATATATGATCGGCTCATGAGACAATAACCCCTGATAATG 2127
 Db 9211 CTTCAATTAATTTGAAAAGAGAGATGAGTAATCAACATTTCCGCTGCTGCTTAT 9270
 Cp 2126 CTTCAATTAATTTGAAAAGAGAGATGAGTAATCAACATTTCCGCTGCTGCTTAT 2067
 Db 9271 CCGTTTTTGGCGATTTTGGCTTCTCTTTTTCCTACCCAGAAACCGTGTGAAGTA 9330
 Cp 2066 CCGTTTTTGGCGATTTTGGCTTCTCTTTTTCCTACCCAGAAACCGTGTGAAGTA 2007
 Db 9331 AAAGATGCTGAAAGATCGATTTGGGTGCAGAGTGGTTTACATCGAATCGGATCTCAACAGC 9390
 Cp 2006 AAAGATGCTGAAAGATCGATTTGGGTGCAGAGTGGTTTACATCGAATCGGATCTCAACAGC 1947
 Db 9391 GGTAAAGATCTTTGAGAGTTTTCGCCCCGGAAGAGCTTTTCCAATGATGAGCAGCTTTTAA 1887
 Cp 1946 GGTAAAGATCTTTGAGAGTTTTCGCCCCGGAAGAGCTTTTCCAATGATGAGCAGCTTTTAA 1887
 Db 9451 GTTCTGCTATGTGGCGGCTATTTATCCCGTATTTGACGCGCGGCAAGAGCAACTCGGTCGC 9510
 Cp 1886 GTTCTGCTATGTGGCGGCTATTTATCCCGTATTTGACGCGCGGCAAGAGCAACTCGGTCGC 1827
 Db 9511 CGCATACACTATTTCTCAGAAATGACTTTGTTGAGTACTCAACAGTCACAGAAAGCACTTT 9570
 Cp 1826 CGCATACACTATTTCTCAGAAATGACTTTGTTGAGTACTCAACAGTCACAGAAAGCACTTT 1797
 Db 9571 ACGGATGGCATGACATGAGTAAGAAATTTATCGAGTGTGCTGATTAACCAAGATGATAACT 9630
 Cp 1766 ACGGATGGCATGACATGAGTAAGAAATTTATCGAGTGTGCTGATTAACCAAGATGATAACT 1707
 Db 9631 CGCGCAACTTACTTCTGACAAAGATCGGAGGACCGAAGAGCTTAACCGCTTTTTCGAC 9690
 Cp 1706 CGCGCAACTTACTTCTGACAAAGATCGGAGGACCGAAGAGCTTAACCGCTTTTTCGAC 1647
 Db 9691 AACATGGGGATCAATGATGACTGCTGATCGTGGGAGCGGAGCTGAATGAAGCCATA 9750
 Cp 1646 AACATGGGGATCAATGATGACTGCTGATCGTGGGAGCGGAGCTGAATGAAGCCATA 1587
 Db 9751 CCAAAACGAGAGCGTGCACACCGATCGCTGTAGCAATGGCAACACGTTGCGCAACTA 9810
 Cp 1586 CCAAAACGAGAGCGTGCACACCGATCGCTGTAGCAATGGCAACACGTTGCGCAACTA 1527
 Db 9811 TTAAGTGGGAACTACTTACTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 9870
 Cp 1526 TTAAGTGGGAACTACTTACTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1467
 Db 9871 GATAAAGTTGACGAGCAGCTTCTGCGCTCGGCGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGAT 9930
 Cp 1466 GATAAAGTTGACGAGCAGCTTCTGCGCTCGGCGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGAT 1407
 Db 9931 AAATCTGGAGCGGTGAGCGGTGGTCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCTGCT 9990
 Cp 1406 AAATCTGGAGCGGTGAGCGGTGGTCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCTGCT 1347
 Db 9991 AAGCGCTCCGCTATCGTACTTATCTACAGAGCGGAGTCAAGCACTATGATGATGATGATGATGATGATGAT 10050
 Cp 1346 AAGCGCTCCGCTATCGTACTTATCTACAGAGCGGAGTCAAGCACTATGATGATGATGATGATGATGATGAT 1287
 Db 10051 AATAGACAGATCGCTGAGATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 10110
 Cp 1286 AATAGACAGATCGCTGAGATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1227
 Db 10111 GTTACTCATATATACATTTAGATTTGATTTTAAACCTTCATTTTAAATTTAAAGNCTAG 10170
 Cp 1226 GTTACTCATATATACATTTAGATTTGATTTTAAACCTTCATTTTAAATTTAAAGNCTAG 1167
 Db 10171 GTGAAGATCGCTTTTGTATATCTCATGACCAATCCCTTTAAGCTGAGTTTTCGTTCCAC 10230
 Cp 1166 GTGAAGATCGCTTTTGTATATCTCATGACCAATCCCTTTAAGCTGAGTTTTCGTTCCAC 1107

QY 1996 ATGTAAACCACTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 2025
 DB 2300 GGTGTGACAAAACACAGAAAGAAAAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2359
 QY 2026 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2085
 DB 2400 GTTGAATACATCATATCT 2419
 QY 2386 TTTTGATATCTATACT 2145
 DB 2420 GTCATGAGCGGATACATATTTTAAATATTAAATATTAAATATTAAATATTAAATATTAAATATT 2479
 QY 2446 TTTTAT 2205
 DB 2480 AATATTGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCA 2539
 QY 2466 AATATTGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCAAAAAGCGGCA 2265
 DB 2540 TAAATTTTGTAAATATATCT 2509
 QY 2566 TAAATTTTGTAAAT 2325
 DB 2600 ATAAATCAAAAT 2659
 QY 2426 ATAAATCAAAAT 2385
 DB 2660 CATATTAAAGACGTGAT 2719
 QY 2486 CATATTAAAGACGTGAT 2445
 DB 2720 TAAATTTTGTAAATATATCT 2779
 QY 2446 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2505
 DB 2780 TAAATTTTGTAAATATATCT 2839
 QY 2506 TAAATTTTGTAAATATATCT 2565
 DB 2840 TGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2899
 QY 2566 TGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2625
 DB 2900 TGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2959
 QY 2626 TGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2685
 DB 2960 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 3019
 QY 2606 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2745
 DB 3020 TATTAAGCGGATGCGGTAAG 3079
 QY 2746 TATTAAGCGGATGCGGTAAG 2805
 DB 3080 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 3124
 QY 2806 GGTGTGACAAAACACAGAAATTAAGAGCGGCAAAAAGAGAAATTAAGAGCGGACACAGCAAA 2850

RESULT 14
 LOCUS SYR72180 2860 bp DNA SYN 04-NOV-1994
 DEFINITION Cloning vector plasmid pTZ180 ampicillin resistance (Amp) and beta-galactosidase complementation (lacZ') genes.
 ACCESSION U37452
 NID 4567984
 KEYWORDS ampicillin resistance; beta-galactosidase-complementation protein; cloning vector (sub-species Cloning vector plasmid pTZ180) DNA.
 SOURCE
 ORGANISM
 UNIDENTIFIED CLONING VECTOR
 ARTIFICIAL SEQUENCE; cloning vectors.
 REFERENCES
 1. (bases 1 to 2860)
 Mead,D.A., Szczesna-Skorupa,E. and Kemper,B.
 TITLE Single-stranded DNA 'Blue' T7 promoter plasmids: a versatile tandem

JOURNAL Promoter system for cloning and protein engineering
 MEDLINE Protein Eng. 1 (1): 67-74 (1996)
 FEATURES 89184389
 source Location/Qualifiers
 1..2860
 /organism="unidentified cloning vector"
 /sub-species="cloning vector plasmid pTZ180"
 /db_xref="taxon:45196"
 217..690
 /gene="LacZ"
 217..690
 /gene="LacZ"
 /codon_start=1
 /transl_table=11
 /function="blue/white selection"
 /product="beta-galactosidase-complementation protein"
 /db_xref="pib:q567890"
 /translation="MTMTNLITETIONSSSSVIRGDEESDTHASLALAVVLAQKPEW
 NPGVTQINFLAAHPHFAKSPNSEAFEPFSEPLATFACTPQALNGLPTGEPFALPHLLPKK11"
 SVTATLASALAPAPFAEPFSEPLATFACTPQALNGLPTGEPFALPHLLPKK11"
 236..352
 /gene="LacZ"
 /note="T7 promoter"
 255..311
 /gene="LacZ"
 /note="MCS"
 /function="multiple cloning site"
 473..942
 /standard_name="fl ori"
 /note="fl origin of replication"
 1960..1920
 /gene="Amp"
 1060..1920
 /gene="Amp"
 /codon_start=1
 /transl_table=11
 /function="ampicillin resistance"
 /product="beta-lactamase"
 /db_xref="pib:q567891"
 /translation="MSIQHPVALIDPEAATFLVLAHPETLVKVKDAEDQLGVVY
 IELQNSKILESIRFELRIPMSIKVLGGZAVLRUAGQJEDVRFYNSYSLIVE
 YSPYKRIIDGTIVRELSNAITMSISNTAANLLITIGFRIELTAFINIMSDIHTFQ
 DRWPELNEATIPNEDDTMPVAMATLEKLTGELLEHLASPGQLDWMWLAIRKVGPI
 LRSALPAGVEATLKSAGERSGRTIAALGHPVKSRIVVITVTSQATMDPEEPGLA
 EIGASLIKHW"
 BASE COUNT 701 a 718 c 709 g 742 t
 ORIGIN
 Query Match 75.3%, Score 2524, BP 42, Expect 2860
 Best local Similarity 99.8%, Prod. No. 0.000000
 Matches 2543; Conservative 0; Mismatches 4; Indels 1; Gaps 0
 DB 314 CACTGACGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 374
 CP 2844 CATTGTGCTGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 2900
 DB 374 GACTTGACACATACGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 400
 CP 2784 GACTTGACACATACGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 2740
 DB 434 GCTTTTCCGCAAAAG 490
 CP 2724 GCTTTTCCGCAAAAG 2600
 DB 494 TAAGCGGCGGCTGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 500
 CP 2664 TAAGCGGCGGCTGCTGCTGCTAATCAATGATCTTCAGCATCTTTTACTTTCACAGAGTCTTCT 2600
 DB 554 GCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
 CP 2604 GCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
 DB 614 AAGCTCTAAATGCGCGAGATCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 670

[illegible]

Cp 1766 ACAGATGAGATGACACTAAGCAATATACAGTCTGTCATTAACCATGATGATACACT 1707
 Db 1456 GGGGCAAAATTAATATACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1515
 Cp 1766 GGGGCAAAATTAATATACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1647
 Db 1516 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1575
 Cp 1646 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1587
 Db 1576 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1635
 Cp 1586 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1527
 Db 1636 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1695
 Cp 1526 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1467
 Db 1695 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1755
 Cp 1466 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1407
 Db 1755 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1815
 Cp 1406 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1347
 Db 1815 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1875
 Cp 1346 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1287
 Db 1875 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1935
 Cp 1286 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1227
 Db 1935 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1995
 Cp 1226 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1167
 Db 1995 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2055
 Cp 1166 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1107
 Db 2055 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2115
 Cp 1106 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1047
 Db 2115 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2175
 Cp 1046 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 987
 Db 2175 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2235
 Cp 986 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 927
 Db 2235 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2295
 Cp 926 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 867
 Db 2295 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2355
 Cp 866 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 807
 Db 2355 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2415
 Cp 806 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 747
 Db 2415 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2475
 Cp 746 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 687
 Db 2475 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2535
 Cp 686 AATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 627

Db 2535 GTAGGCGGAGGCTCGAAGACAG 2595
 Cp 526 GAAAGCGGAG 567
 Db 2596 TATCTTTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2655
 Cp 566 TATCTTTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
 Db 2656 TCGTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2715
 Cp 586 TCGTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447
 Db 2716 GCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2775
 Cp 446 GCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
 Db 2776 AACCGTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2835
 Cp 386 AACCGTATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
 Db 2836 AGCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2895
 Cp 326 AGCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 327

RESULT 15
 ID E11270 standard: DNA, UNC: 2927 BP.
 AC E11270;
 NI d1109607
 DT 08-OCT-1997 (Rel. 52, Created)
 DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)
 DE Nucleotide sequence of pTRP.
 KW 199610379-A/2
 OS unclassified
 OC unclassified
 RN [1]
 RP 1-2927
 RA Ishizuka T., Kihira Y., Tanaka T., Matsuo T.;
 RA "PRODUCTION OF RECOMBINED HUMAN MYOGLOBIN";
 RL Patent number JP 199610379-A/2, 23-APP-1996.
 PL ORIENTAL YEAST CO LTD.
 OS None
 CC OC Artificial sequences.
 CC OC JP 199610379-A/2
 CC PN 23-APP-1996
 CC PD 07-OCT-1994 JP 1994268121
 CC PI ISHIZUKA TOSHIHIRO, KIHIRA YASUNORI, TANAKA TOSHIO,
 CC PI MATSUO TAKESHI
 CC PC G12N15/09,C07H21/04,C07K14/805,G12P21/02,G12P21/02,G12P1-19;
 CC CC strandedness: Double;
 CC CC topology: Circular;
 CC CC key Location/Qualifiers
 CC FT source 1..2927
 CC FT /organism="Artificial sequences"
 CC FT promoter 267..326
 CC FT /note="trp promoter derived from E. coli K12"
 CC FT CDS 1127..1197
 CC FT /product="beta-lactamase"
 CC FT Location/Qualifiers
 FH FT source 1..2927
 FH FT /organism="unidentified"
 FT Sequence 2927 BP; 725 A; 736 C; 718 G; 748 T; 0 other;
 SQ

Query Match 75.2%, Score 2529, DB 10; Length 2927.
 Best Local Similarity 59.7%, Pred. No. 0.00-0.00;
 Matches 2541; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
 Db 380 TCAGTGGCGGCTGCTTTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
 Cp 2845 TCAGTGGCGGCTGCTTTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2786

US-09-020-716-2.rge

Wed Mar 17 09:43:27 1999

```

Cp 526 3TAAAGCCGACGCTCGGAAAGAGAGAGCTATAGAGGAGCTTCCAGGGGGAACCCCTGG 567
Db 2680 TATCTTTATAGTCTGTGCGGTTTCGCCAGCTCTGACTTGAGCGTCGATTTTGTGATGC 2719
Cp 566 TATCTTTATAGTCTGTGCGGTTTCGCCAGCTCTGACTTGAGCGTCGATTTTGTGATGC 507
Db 2720 TCGTCAGGGGGGCGGAGCCTATGGAAGAACGCCAGCAAGCGGCCCTTTTACGGTTCCCTG 2779
Cp 506 TCGTCAGGGGGGCGGAGCCTATGGAAGAACGCCAGCAAGCGGCCCTTTTACGGTTCCCTG 447
Db 2780 GCGTTTGGTGGGCTTTTGGTCACATGTTCTTTCTCGGTTATCCGCTGATTTCTGTGGAT 2839
Cp 446 GCGTTTGGTGGGCTTTTGGTCACATGTTCTTTCTCGGTTATCCGCTGATTTCTGTGGAT 387
Db 2840 AACGCTATTACGCGCTTTTGGTACGCTGATACCGCTGCGCGGAGCGGAGCGCGAGCGC 2899
Cp 386 AACGCTATTACGCGCTTTTGGTACGCTGATACCGCTGCGCGGAGCGGAGCGCGAGCGC 327
Db 2900 AACGCTATTACGCGCTTTTGGTACGCTGATACCGCTGCGCGGAGCGGAGCGCGAGCGC 2927
Cp 326 AACGCTATTACGCGCTTTTGGTACGCTGATACCGCTGCGCGGAGCGGAGCGCGAGCGC 299

```

Search completed: Sun Mar 14 13:14:36 1999
Job time : 8281 secs.

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

WORLD (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPsrch_en 0 1 - n a database search, using Smith-Waterman algorithm
Run on: Sun Mar 14 15:24:16 1999: Maspar time 485.47 Seconds
943.032 Million cell updates/sec

Tabular output not generated.

Title: >US-09-020-716-2
Description: (1:3365) from US09020716 seq
Perfect score: 3365
N.A. Sequence: 1 TCGACCTCGAGGGGGGGG
Comp: AGCTGAGTTCGCGCGG

Scoring table: TABLE default.
Gap 5

Nmatch STD : Dbase 0: Query 0
Searched: 189442 seqs, 58026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 10.017: Variance 5.937: scale 1.687
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	ID	Description	Pred. No.	
1	4855	84.8	534	25	T43137	PUMIGIT sequence incl	0.00e+00
2	4844	83.9	727	19	V28910	Plasmid pWP3169 enco	0.00e+00
3	2820	93.8	5585	14	V82418	Plasmid pRP6-IL2	0.00e+00
4	2818	83.7	5585	29	T61430	Interleukin-2 express	0.00e+00
5	2774	82.4	5924	6	Q10650	K lactis/S. cerevisae	0.00e+00
6	2773	75.2	2927	24	T31789	Plasmid pTRP	0.00e+00
7	2772	75.2	2927	24	T31789	Plasmid pTRP	0.00e+00
8	2753	71.2	5243	24	T41898	Vector pET Trc S05/N1	0.00e+00
9	2749	71.2	5248	17	T08489	Plasmid pET Trc S05/N	0.00e+00
10	2740	70.7	3554	11	T22275	Plasmid pASK46 for ex	0.00e+00
11	2735	70.5	6714	13	Q79545	pW3254b Plasmid expre	0.00e+00
12	2734	67.0	3812	11	G24276	Plasmid pASK60-Strep	0.00e+00
13	2754	67.0	4515	8	Q46823	plISC-SE vector.	0.00e+00

C	14	2254	67.0	4640	5	Q28910	RBI, PDI coexpression	0.00e+00
C	15	2232	65.0	5178	27	T49876	Pret-Splice	0.00e+00
C	16	2230	65.0	6205	27	T49877	Autoregulatory vector	0.00e+00
C	17	2232	66.0	7474	31	T71320	Plasmid pTb73 encodin	0.00e+00
C	18	2215	65.8	6414	3	Q20733	PAD-CMV1 expression v	0.00e+00
C	19	2215	65.8	6414	1	Q36283	PAD-CMV1 contig a Tum	0.00e+00
C	20	2203	65.5	3681	2	Q13578	Plasmid pKSEL5	0.00e+00
C	21	2205	65.5	5314	2	Q13576	Plasmid pMTV1 contain	0.00e+00
C	22	2203	65.5	6414	3	Q20766	PAD-CMV1	0.00e+00
C	23	2202	65.4	3649	39	V14340	Plasmid pBSGFP expres	0.00e+00
C	24	2202	65.4	4145	7	Q40279	Sequence of clone pS2	0.00e+00
C	25	2200	65.4	4539	14	Q87347	Plasmid pIN1	0.00e+00
C	26	2202	65.4	4792	15	Q84696	Plasmid pS contg. N.	0.00e+00
C	27	2202	65.4	5042	15	Q84694	Plasmid glucosylase	0.00e+00
C	28	2202	65.4	5356	25	T43794	Plasmid pPRIPAT (rat	0.00e+00
C	29	2200	65.4	10930	20	Q81226	Plasmid pM16-1	0.00e+00
C	30	2200	65.4	10950	20	Q81225	Plasmid pM16	0.00e+00
C	31	2198	65.3	4277	7	Q40280	Sequence of clone pS2	0.00e+00
C	32	2198	65.3	6926	7	Q40281	Sequence of clone pS2	0.00e+00
C	33	2195	65.2	14311	24	T38744	Plasmid pSVFL(-) inc	0.00e+00
C	34	2187	65.0	7238	9	G50201	Vector p37M1-10C	0.00e+00
C	35	2170	64.5	4118	35	T69189	Construct pGEM-nrk(Ka	0.00e+00
C	36	2170	64.5	4118	35	T69188	Construct pGEM-nrk(Ka	0.00e+00
C	37	2171	64.5	6295	39	V02043	Plasmid pWRG3195 enco	0.00e+00
C	38	2166	64.4	4540	32	T69891	Human cyclin D1-human	0.00e+00
C	39	2166	64.4	4824	20	T08975	Shortened C1 gene in	0.00e+00
C	40	2167	64.4	4883	4	Q25709	pDEL10	0.00e+00
C	41	2167	64.4	4883	8	Q51192	pDEL110; plasmid DNA	0.00e+00
C	42	2167	64.4	5620	4	Q27489	Plasmid pVE108 used 1	0.00e+00
C	43	2167	64.4	5620	4	Q27489	pVE108	0.00e+00
C	44	2167	64.4	5642	8	Q51193	pDE4; plasmid DNA rep	0.00e+00
C	45	2167	64.4	7050	7	Q40419	Sequence of pTE4.	0.00e+00

ALIGNMENTS

RESULT 1
T43137 standard; DNA: 5534 BP.
AC T43137:
DE 10-FEB-1997 (first entry)
DE PUMIGIT sequence including upstream activating sequence.
KW Transgenic plant, gene expression, upstream activating sequence;
KW UAS: transactivating protein, Gal4; herbicide resistance,
KW polyhydroxybutyrate; safety, pUMIGIT; ds.
OS Synthetic.
PN CA2150039-A.
PD 09-AUG-1996.
PF 24-MAY-1995; 150039.
PR 08-FEB-1995; GS-002456.
PR (UYWA-) UNIV WARWICK.
PA Bennett M, May S, Ramsay N:
PI WPI; 96-486150/49.
DR Control of genes in transgenic plants using an upstream activating
PT sequence activated by a transactivating protein expressed using a
PT separate promoter
PT
PS Example; Fig 5F: 48pp: English.
PS Reporter plasmid pUMIGIT (T43137), or pUAS Minimal promoter Gus
PS Init terminator, contains a beta-glucuronidase (GUS) reporter
PS gene under the control of the 46S cauliflower-mosaic virus minimal
PS promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream
PS activating sequence (UAS) of pUMIGIT is activatable by yeast
PS transactivating protein GAL4. In a novel method for controlling
PS gene expression, a first transgenic plant carrying a gene encoding
PS a desired phenotype (herbicide resistance of polyhydroxybutyrate
PS prodn.) operatively linked to a UAS recognition site is pollinated
PS by a second transgenic plant carrying Gal4 DNA (see also T43136).
PS The transgene is fully expressed in F1 hybrid plants but sequestrates
PS apart in subsequent generations, improving environmental safety.
SQ Sequence 5534 BP. 1400 A. 1355 C. 1487 G. 1363 T.

Query Match 84.8% Score 2855.7E 25. Length 5534;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2855. Conservative 0. Mismatches 0. Indels 0. Gaps 0.

[illegible]

```

PT creating animal models for drug screening
PS Example 1: Fig 9b-g: 82pp; English.
CC Vector pTet-Splice (T49876) is utilized in the construction of
CC autoregulatory vector plasmid pTet-trak (T49877). Splice-PA was
CC made by ligating the SV40 small T antigen intervening sequence and
CC the SV40 early polyA sequence into pBSK11+. The XhoI-SalI fragment
CC of pHC13-3 (contg. 7 copies of the tet operator upstream of minimal
CC promoter Tetp) was cloned upstream of the splice/polyA sequence of
CC pTet-Splice. A modified tetracycline
CC transactivator (tTAK) gene was cloned into pTet-Splice to form pTet-
CC transactivator (tTAK) gene. The construct provides rapid, reversible control of gene
CC expression in eukaryotic cells or transgenic animals, e.g. for
CC heterologous protein, RNA or antisense sequence prodn., or for
CC creating animal models of human disease.
CC Sequence 5178 BP: 1383 A; 1169 C; 1240 G; 1386 T;
SQ
Query Match 66.0%, Score 22.2, EB 27, Length 5178;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2957 ctcagggggggcccggtaccacgcttttgccttttagtgaggggttaattgcgcgtt 3016
QY 6 CTCAGAGGGGGGGCCCGGTACCCACCTTTTGTCCCTTATGAGGGGTAAATTGCGCGCTT 65
Db 3017 ggcgttaattcgttcagctgtttctctgtgtgaaattttatccgctcacattccaca 3076
QY 66 GCGTAAATCATGGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCACA 125
Db 3077 caactacagccggaagcataaagataaagctggggtgcttaagtgagtaagtaact 3136
QY 126 CAACATACAGCCGGAACATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAGTAACT 185
Db 3137 cacattaattgcttgcctcactgcccgcctttccagtcgggaaacactgtctgcgcagt 3196
QY 186 CACATTAATTGGTTTGGCTGCTCACTGCTGCTTTCACGTGCGGAAAGCTGTCTGCTGCT 245
Db 3197 gatttaataatccgcaaacgagcgaggagagcggtttcgtattgagcgctcttccgc 3256
QY 246 GCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGGTATTTGGTGGCTTTCGCG 305
Db 3257 ttctcgcctcactgactgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 3316
QY 306 TTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 335
Db 3317 ctcaaggggttaatacgtttatccacagaaatcagggtataacgaggaagaaacatgtg 3376
QY 366 CTCAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGTAACGCAGAAAGAAACATGTG 425
Db 3377 agcaaaaggccagcaaaaggccaggaacccgtataaaaggccgcttgcgttttttcca 3436
QY 426 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCTTTCCTGGCGTTTTCCTCA 485
Db 3437 taggtccgccccctgagcagcatcacaaaaatcgacctcaagtcagagtgagtgcaaaa 3496
QY 486 TAGGTCCGCCCCCTGACGAGCATCACAAAAAAGCAAGTCAATCAGAGGTGGGAAA 545
Db 3497 ccgacaggaactataagataccagagcttttccctctgagactccctctgagcgtctcc 3556
QY 546 CCGGACAGGACTATAAAGATACAGAGGCTTTCGCGCTGGAAGCTCCCTGTCGCGCTCCG 605
Db 3557 tcttcgacctgcgcttaccgataacctgtcgcctttctcccttcgggagagcgtggc 3616
QY 606 TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAGGTGGC 665
Db 3617 gctttctcagctcagcgtgtaggtatctcagttcagttcagttcagttcagttcagttc 3676
QY 666 GCTTCTCATAGCTCACGCTTAGGTATCTCAGTTCGTTAGGTGTTAGGTGTTAGGTGTT 725
Db 3677 gggctgtgtgacgaaccccccggttcagcccgacgctgagccttatccgtataactatcc 3736
QY 726 GGGCTGTGTGACGAACCCCCCGCTTCACCCCGCAAGCTTTCGCGCTTATCCGGTAACTAG 785
Db 3737 tottgagtcacaccccggttaagacagcattatccgactatccgactatccgactatcc 3796

```

```

RESULT 15
ID T49876 standard; DNA; 5178 BP.
AC T49876;
DT 28-MAR-1997 (first entry)
DE pTet-Splice.
KW pTet-Splice; pTet-tTAK; tetracycline transactivator; promoter;
KW gene expression; transgenic animal; animal model; drug screening;
KW vector; ds
OS Synthetic.
FH Key Location/Qualifiers
FT misc_rna complement (2591)
FT FT /*tag= a
FT FT /note= "putative start site of trxn"
FT FT complement (2616..2622)
FT FT /*tag= b
FT tata_signal
FT WO640946-A1
PD 19-DEC-1996
PR 07-JUN-1996: 010109
PR 07-JUN-1995: US-474169.
PA (UYTA ) UNIV YALE.
PI Schatz DG;
DE WP1 97-07273/07
PI Nucleic acid encoding tetracycline transactivator fusion protein.
PI provides rapid and reversible control of gene expression, e.g. for
PI

```


REFERENCE
1 (bases 1 to 721)
Tovoda A., Shiba T. and Sakaki, Y.

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, I. and Sakakima,
I.

TITLE Human sapiens genomic DNA, chromosome 21q

JOURNAL
Homo sapiens genomic DNA, chromosome 11
published Only in DataBase (1998) In press

REFERENCE
2 (bases 1 to 721)

AUTHORS
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
Direct Submission

TITLE Direct Submission
JOURNAL Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank database

Submitted (29 Oct 1997) to the
Masahira Hattori, Kitasato University,
Department of Sc

Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228,
Tel. 0427-78-9732

(E-mail: hattori@hg.ims.u-tokyo.ac.jp, Tel: 0427-78-5737)
Fax: 0427-78-9561)

FEATURES Location/Qualifiers

FEATURES	LOCATION/QUALITY
SOURCE	1. 721

Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@ims.kit.ac.jp, Tel:0427-78-9742,
Fax:0427-78-9561)

FEATURES
SOURCE
Location/Qualifiers
1..696
Organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="F84SfN15"
/map="21q"

BASE COUNT 189 a 153 c 171 q 175 t 10 others

ORIGIN
Query Match 18.5% Score 624; DB 27; Length 696;
Best Local Similarity 99.1% Pred No. 0.00e+00;
Matches 660; Conservative 0; Mismatches 0; Indels 5; Gaps 5;
Db 31 GTTACGATGGCATACATTAAGTAAGTAATATGATGTTGGTATAACCATGAGTGATAAC 90
Cp 1769 GTTACGATGGCATACATTAAGTAAGTAATATGATGTTGGTATAACCATGAGTGATAAC 1710
Db 91 ACTGCGGTAACATTAATTTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 150
Cp 1739 ACTGCGGTAACATTAATTTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1650
Db 151 CAAACATGGGATGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 210
Cp 1649 CAAACATGGGATGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1590
Db 211 ATACCAAAAGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 270
Cp 1589 ATACCAAAAGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530
Db 271 TATTAACTGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330
Cp 1524 TATTAACTGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
Db 331 GTCGATAAACTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 390
Cp 1469 GTCGATAAACTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1410
Db 441 GATAAATCTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
Cp 1409 GATAAATCTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
Db 451 GGTAAAGGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 509
Cp 1349 GGTAAAGGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
Db 520 GAAATAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 569
Cp 1289 GAAATAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
Db 570 GAAAGTTATCT-TATATATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
Cp 1230 GAAAGTTATCT-TATATATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
Db 629 CTATGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Cp 1170 CTATGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
Db 688 TCACTGA 695
Cp 1110 TCACTGA 1104

RESIDUE AC009765 646 bp DNA GSS 05-JUL-1998
FEATURES
Homo sapiens genomic DNA, 21q region, clone: F84SfN15, genomic
survey sequence.
AC009765
NID G3289751

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBLISHED ONLY IN DATABASE (1998) In press
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
2 (Bases 1 to 696)
Direct Submission
Submitted (06-09-1998) to the EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@ims.kit.ac.jp, Tel:0427-78-9742,
Fax:0427-78-9561)

FEATURES
source
Location/Qualifiers
1..696
Organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="F84SfN15"
/map="21q"

BASE COUNT 183 a 165 c 185 q 162 t 1 others

ORIGIN
Query Match 18.5% Score 624; DB 27; Length 696;
Best Local Similarity 99.1% Pred No. 0.00e+00;
Matches 660; Conservative 0; Mismatches 0; Indels 5; Gaps 5;
Db 31 GTTTTGGTCAGCAGCAAAAGCTGGTAAATTAATAATGATGATGATGATGATGATGATGATGATGAT 90
Cp 2039 GTTTTGGTCAGCAGCAAAAGCTGGTAAATTAATAATGATGATGATGATGATGATGATGATGATGAT 1989
Db 91 CGAGTGGGTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150
Cp 1979 CGAGTGGGTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 151 GAAGACGCTTTGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 210
Cp 1919 GAAGACGCTTTGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1869
Db 211 GTATTGAGGCGGCGGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 270
Cp 1859 GTATTGAGGCGGCGGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1809
Db 271 GTTGAT 330
Cp 1799 GTTGAT 1749
Db 331 TGCAGTGTGCTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
Cp 1739 TGCAGTGTGCTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1689
Db 391 GAGACGCGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
Cp 1679 GAGACGCGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629
Db 451 CATCGTGTGCTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
Cp 1619 CATCGTGTGCTATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
Db 511 GGCCTGTAGCAATGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Cp 1560 GGCCTGTAGCAATGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1510
Db 571 GTTCTGGGTAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
Cp 1501 GTTCTGGGTAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1451
Db 631 GATCTGGGTAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
Cp 1451 GATCTGGGTAACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401

US-09-020-716-2.1st

Wed Mar 17 09:43:32 1999

```

Cp 1442 CGCTGGGCGCTT-CGGGCTGG-CTGGTTTATTGCTGATAAAATCTGGAGCGGTGACCGTG 1385
Db 690 GGTCTC 695
Cp 1384 GGTCTC 1379

RESULT 10 AG003576 696 bp DNA GSS 19-DEC-1997
LOCUS Homo sapiens genomic DNA, 21q region, clone: P864SpN15. genomic
DEFINITION survey sequence.
ACCESSION AG003576
NID g2705702
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P864SpN15.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 696)
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
Homo sapiens genomic DNA, chromosome 21q
Published Only in Database (1997) In press
REFERENCE 2 (bases 1 to 696)
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
Direct Submission
Submitted (19-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagami-hara 228, Japan
(E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
FEATURES
Location/Qualifiers
source 1..696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P864SpN15"
/map="21q"
BASE COUNT 183 a 165 c 185 g 162 t 1 others
ORIGIN
Query Match 18.5%; Score 624; DB 26; Length 696;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 650; Conservative 0; Mismatches 0; Indels 6; Gaps 6;

Db 31 GTTTTGCTACCCAGAACCGCTGGTGAAGTAAAGATGCTCAAGATCAGTTGGTGCA 90
Cp 2039 GTTTTGCTACCCAGAACCGCTGGTGAAGTAAAGATGCTCAAGATCAGTTGGTGCA 1980
Db 91 CGAGTGGGTATGATGAGTCAAGTCAAGCGGTGAAGATCCTTGAGAGTTTCGCCCC 150
Cp 1979 CGAGTGGGTATGATGAGTCAAGTCAAGCGGTGAAGATCCTTGAGAGTTTCGCCCC 1920
Db 151 GAAGACGTTTTTCAATGATGAGTCAAGTCAAGCGGTGAAGATCCTTGAGAGTTTCGCCCC 210
Cp 1919 GAAGACGTTTTTCAATGATGAGTCAAGTCAAGCGGTGAAGATCCTTGAGAGTTTCGCCCC 1860
Db 211 CGTATTGACCGCGGCAAGAGAACTTCGGTTCGGCGCATACACTATTCAGATGACTTG 270
Cp 1859 GCTATTGACCGCGGCAAGAGAACTTCGGTTCGGCGCATACACTATTCAGATGACTTG 1800
Db 271 GTTGAGTACTACCAAGTCAAGAGAACTTCGGTTCGGCGCATACACTATTCAGATGACTTG 330
Cp 1799 GTTGAGTACTACCAAGTCAAGAGAACTTCGGTTCGGCGCATACACTATTCAGATGACTTG 1740
Db 331 TGCATGTGTCGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 390
Cp 1739 TGCATGTGTCGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 391 GGAGGACCGCAAGAGAGTAAACCGCTTTTTCGACAACTGGGGGATCATGTAACTCGCGTT 450
Cp 1679 GGAGGACCGCAAGAGAGTAAACCGCTTTTTCGACAACTGGGGGATCATGTAACTCGCGTT 1620
Db 451 GATCGTTGGGACCGGAGCTGAATGAAGGCAATACCAAGCAGCGGTGAACACCAAGAT 510

```

```

Cp 1619 GATCGTTGGGAACCGGASCTGAATGAAGCCATACCAACGAGAGCTGGA-CACACGAT 1561
Db 511 GGCCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTGGCAAACTACTACTACTAG 570
Cp 1560 G-CCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTGGCAAACTACTACTACTAG 1502
Db 571 CTTCGCGGCAACAAATTAATAGACTGGATGAGCGGATAAAGTTGCGAGGACACTTCIG 630
Cp 1501 CTTCGCGG-CACAATTAATAGACTGGATGAGCGGATAAAGTTGCGAGGACACTTCIG 1443
Db 631 CGCTCGGCCCTTCGCGCTGGCTGGTTATTGCTGATAAACTGAGCGCGGTGACG-TG 689
Cp 1442 CGCTCGGCCCTT-CCGGCTGG-CTGGTTATTGCTGATAAACTGAGCGCGGTGACG-TG 1385
Db 690 GGTCTC 695
Cp 1384 GGTCTC 1379

RESULT 11 AG013858 720 bp DNA GSS 10-SEP-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: 762015SpN14. genomic
DEFINITION survey sequence.
ACCESSION AG013858
NID g3560348
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 762015SpN14.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 720)
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
Homo sapiens genomic DNA, chromosome 21q
Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 720)
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
Direct Submission
Submitted (10-SEP-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagami-hara 228, Japan
(E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
FEATURES
Location/Qualifiers
source 1..720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="762015SpN14"
/map="21q"
BASE COUNT 172 a 178 c 170 g 186 t 14 others
ORIGIN
Query Match 18.5%; Score 624; DB 28; Length 720;
Best Local Similarity 97.2%; Pred. No. 0.00e+00;
Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTCCGCTTACCGGATACCTGCGCCTT 94
Cp 586 AGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTCCGCTTACCGGATACCTGCGCCTT 645
Db 95 CTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTG 154
Cp 646 CTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGTG 705
Db 155 TAGTGTGCTGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTCAGCGGACGCTGCT 214
Cp 706 TAGTGTGCTGCTCCAGCTGGGTGTGTGTCAGCAACCCCGCTTCAGCGGACGCTGCT 765
Db 215 GCCTTATCCCGTAACTATCGTCTTGGATCCAAACCCCGGTGAACACGACGCTTATCG 274
Cp 746 GCCTTATCCCGTAACTATCGTCTTGGATCCAAACCCCGGTGAACACGACGCTTATCG 825

```


Wed Mar 17 09:43:32 1998

```

Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel. 301 838 0200
Fax: 301 838 0208
E-mail: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES             source
    1..644
        /organism="Homo sapiens"
        /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
        HindIII"
        /db_xref="taxon:9606"
        /clone="2374C24"
        /clone_lib="Cir-Hsp"
        /sex="Male"
        /cell_type="Sperm"

BASE COUNT      135 a  183 c  145 g  179 t      2 others
ORIGIN
Query Match      18.1%; Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCATGCGCGTCTTTACACAGCTGCTGACCTGGAAATCTCTGGGATTACCCAACTTAAT 64
Cp 2845 TCATGCGCGTCTTTTACACAGCTGCTGACCTGGAAATCTCTGGGATTACCCAACTTAAT 2785
Db 65 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAGAT 124
Cp 2785 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 2725
Db 125 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 183
Cp 2725 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 2666
Db 184 TTAAGCGGGGGGGTGTGGTGTACGCGAGCGTGACCGCTACACTTGCACACGCGCTA 243
Cp 2665 TTAAGCGGGGGGGTGTGGTGTACGCGAGCGTGACCGCTACACTTGCACACGCGCTA 2606
Db 244 GCGCGCGTCTTTTCCGCTTCCTCCCTTCCTTCGCGGAGCTTCGCGGCTTCGCGG 303
Cp 2605 GCGCGCGTCTTTTCCGCTTCCTCCCTTCCTTCGCGGAGCTTCGCGGCTTCGCGG 2546
Db 304 CAAGCTCTAAATCGGGGGTCTCCCTTTAGGGTTCGCACTTACTGCTTTTACGCGCTGAC 363
Cp 2545 CAAGCTCTAAATCGGGGGTCTCCCTTTAGGGTTCGCACTTACTGCTTTTACGCGCTGAC 2486
Db 364 CCCAAATCTTATAGGTGATGTTTACGCTATGAGGCTATCGGCTGATAGAGGCTT 423
Cp 2485 CCCAAATCTTATAGGTGATGTTTACGCTATGAGGCTATCGGCTGATAGAGGCTT 2426
Db 424 TTTGCGGCTTTGAGCTTGGAGTCCAGCTTCTTAAATATATATAGGAGTTTTCGCGATTCG 483
Cp 2425 TTTGCGGCTTTGAGCTTGGAGTCCAGCTTCTTAAATATAGGAGTTTTCGCGATTCG 2366
Db 484 ACAACATCTAAGCTATCTCGGCTATCTCTTTTGTATATAGGAGTTTTCGCGATTCG 543
Cp 2365 ACAACATCTAAGCTATCTCGGCTATCTCTTTTGTATATAGGAGTTTTCGCGATTCG 2306
Db 544 GCTTATGGTTAAATATAGCTGATTTAACAAAAATTTTAAACAAAAATTTTAAACAAAAATA 603
Cp 2305 GCTTATGGTTAAATATAGCTGATTTAACAAAAATTTTAAACAAAAATTTTAAACAAAAATA 2246

```

```

TITLE
JOURNAL
COMMENT

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel. 301 838 0200
Fax: 301 838 0208
E-mail: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES             source
    1..644
        /organism="Homo sapiens"
        /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
        HindIII"
        /db_xref="taxon:9606"
        /clone="2374C24"
        /clone_lib="Cir-Hsp"
        /sex="Male"
        /cell_type="Sperm"

BASE COUNT      135 a  183 c  145 g  179 t      2 others
ORIGIN
Query Match      18.1%; Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCATGCGCGTCTTTACACAGCTGCTGACCTGGAAATCTCTGGGATTACCCAACTTAAT 64
Cp 2845 TCATGCGCGTCTTTTACACAGCTGCTGACCTGGAAATCTCTGGGATTACCCAACTTAAT 2785
Db 65 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 124
Cp 2785 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 2725
Db 125 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 183
Cp 2725 GCGCTTCGAGCACATCCCGCTTCGCCAGCTGCGTAAAGCAAGAGAGAGAGAGAT 2666
Db 184 TTAAGCGGGGGGGTGTGGTGTACGCGAGCGTGACCGCTACACTTGCACACGCGCTA 243
Cp 2665 TTAAGCGGGGGGGTGTGGTGTACGCGAGCGTGACCGCTACACTTGCACACGCGCTA 2606
Db 244 GCGCGCGTCTTTTCCGCTTCCTCCCTTCCTTCGCGGAGCTTCGCGGCTTCGCGG 303
Cp 2605 GCGCGCGTCTTTTCCGCTTCCTCCCTTCCTTCGCGGAGCTTCGCGGCTTCGCGG 2546
Db 304 CAAGCTCTAAATCGGGGGTCTCCCTTTAGGGTTCGCACTTACTGCTTTTACGCGCTGAC 363
Cp 2545 CAAGCTCTAAATCGGGGGTCTCCCTTTAGGGTTCGCACTTACTGCTTTTACGCGCTGAC 2486
Db 364 CCCAAATCTTATAGGTGATGTTTACGCTATGAGGCTATCGGCTGATAGAGGCTT 423
Cp 2485 CCCAAATCTTATAGGTGATGTTTACGCTATGAGGCTATCGGCTGATAGAGGCTT 2426
Db 424 TTTGCGGCTTTGAGCTTGGAGTCCAGCTTCTTAAATATATAGGAGTTTTCGCGATTCG 483
Cp 2425 TTTGCGGCTTTGAGCTTGGAGTCCAGCTTCTTAAATATAGGAGTTTTCGCGATTCG 2366
Db 484 ACAACATCTAAGCTATCTCGGCTATCTCTTTTGTATATAGGAGTTTTCGCGATTCG 543
Cp 2365 ACAACATCTAAGCTATCTCGGCTATCTCTTTTGTATATAGGAGTTTTCGCGATTCG 2306
Db 544 GCTTATGGTTAAATATAGCTGATTTAACAAAAATTTTAAACAAAAATTTTAAACAAAAATA 603
Cp 2305 GCTTATGGTTAAATATAGCTGATTTAACAAAAATTTTAAACAAAAATTTTAAACAAAAATA 2246

```

```

/clone="21"
/clone="c48B2A23"
/map="21g"

BASE COUNT      194 a  171 c  176 g  194 t  12 others
ORIGIN
Query Match      18.4%; Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTGCGTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGC 100
Cp 934 ATCTGCGTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGC 993
Db 101 AAACAAGCAGCGTGTGAGTGGTG- TTTTGTGTTTGAAGCAGCAGATTAACGCGCAGA 159
Cp 994 AAACAAGCAGCGTGTGAGTGGTG- TTTTGTGTTTGAAGCAGCAGATTAACGCGCAGA 1053
Db 150 AAAAAGGATCTCAAGAGATCTTTTATCTTTCTACGGGGTCTGACGCTCAGTGAAC 219
Cp 1054 AAAAAGGATCTCAAGAGATCTTTTATCTTTCTACGGGGTCTGACGCTCAGTGAAC 1113
Db 220 GAAACCTACGTTTAAAGGATTTTGTGATGATATCAAAAAGATCTTCACTAGATC 279
Cp 1114 GAAACCTACGTTTAAAGGATTTTGTGATGATATCAAAAAGATCTTCACTAGATC 1173
Db 280 CTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACCTTGTCT 339
Cp 1174 CTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACCTTGTCT 1233
Db 340 GACATTTACCAATGCTTAATGATGATGATATCAAAAAGATCTTCACTAGATC 399
Cp 1234 GACATTTACCAATGCTTAATGATGATGATATCAAAAAGATCTTCACTAGATC 1293
Db 400 TCCATAGTTCCTGACTCCCGCTGCTGTAGTAACTACGATACGCGAGGCT-ACCATCT 458
Cp 1294 TCCATAGTTCCTGACTCCCGCTGCTGTAGTAACTACGATACGCGAGGCTTACCATCT 1353
Db 454 GAGATTCAGTGTGATGATGATACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
Cp 1354 GAGATTCAGTGTGATGATGATACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1413
Db 519 ATAACACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577
Cp 1414 ATAACACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Db 578 CATCAGTCTATTAATTTGCGCGGAGCTAGAGTAAAGTATTTTCGCCAGTTAATAGT 637
Cp 1473 CATCAGTCTATTAATTTGCGCGGAGCTAGAGTAAAGTAA-GTAGTT-CGCCAGTTAATAGT 1530
Db 638 TTGCGGAGCTTTGTCGATTTGCTGGAGGATGATGCTGACGCGCTCCTTGGTAA 697
Cp 1531 TTGCGGAGCTTTGTCGATTTGCTGAGGATGATGCTGACGCGCTCCTTGGTAA-T 1589
Db 697 TTTTATTTAGCTTGGAGTTCGAGTAAAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 746
Cp 1590 TTTTATTTAGCTTGGAGTTCGAGTAAAGTAAAGGAGAGAGAGAGAGAGAGAGAGAG 1636

RESULT 14
LOCUS      A0115013      644 bp      DNA      GSS      28-AUG-1998
DEFINITION      C11-HSP-2374C24.F C11-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION      A0115013
NID      93491134
KEYWORDS      GSS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 644)
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, F., Wible, C., Shizuya, H., Simon, M. and

```

```

DB 614 TTAACGCTTACCAATTC 614
      TTTTTTTTTTTTTTTT
CY 2245 TTAACGCTTACCAATTC 2245

FEATURES             16
Source              AC040352      612 bp      DNA      100-100-1998
DEFINITION          C17-HSP-2327K21, 19 C17 HSP Homo sapiens genomic clone 2327K21,
                    acc. # AC040352, sequence.
ACCESSION           AC040352
NI                  43496184
KEYWORDS            GSS.
SOURCE              Human.
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Hominidae; Homo.
REFERENCE           1 (bases 1 to 692)
AUTHORS             Adams, M. D., Kornblow, S. D., Field, C. E., Bass, S., Linher, K.,
                    Alden, K., Jett, K., Stauder, J., Sub, E., Wible, C., Shizuya, H.,
                    Simon, M., and Venter, A.
TITLE               Use of a Random HA* End Sequence Database for Sequence Ready Map
                    Building
JOURNAL             Unpublished (1997)
COMMENT             Other_GSS: C17-HSP-2327K21.FP
                    Contact: Mark Adams
                    Department of Eukaryotic Genomics
                    The Institute for Genomic Research
                    4712 Medical Center Dr., Rockville, MD 20850, USA
                    Tel: 401 848 0200
                    Fax: 401 848 0208
                    Email: mdamads@tiara.org
                    Clones are available from Research Genetics (info@resgen.com). BAC
                    end search page:
                    http://www.resgen.com/BAC/bacmap3/bacmap3.cgi?map=C17-HSP-2327K21
                    Seq primer: M13-21
                    Class: BAC ends.
FEATURES             16
Source              1. 692
                    Location/Qualifiers
                    1..692
                    /organism="Homo sapiens"
                    /note="Vector: pBcl-BAC11, Site_1: HindIII; Site_2:
                    HindIII"
                    /db_xref="taxon:9606"
                    /clone="2327K21"
                    /clone_lib="C17-HSP"
                    /sex="Male"
                    /colony_type="Sperm"
BASE COUNT          158 a 134 c 186 g 155 t
ORIGIN

```

Query Match: 17.9% Score 603; DB 27; Length 692;

Best Local Similarity: 99.8%; Prod.No. 0.000000;

Matches 604; Conservatize 0; Mismatches 1; Indels 0; Gaps 0;

```

DB 48 GCTTGGGTAATCATGTCATAGCTGTTCTTCCTGCTGGAAATGTTATCCGTCACAATTC 147
      TTTTTTTTTTTTTTTTTT
CY 62 GTTGGGTAATCATGTCATAGCTGTTCTTCCTGCTGGAAATGTTATCCGTCACAATTC 121
      TTTTTTTTTTTTTTTTTT
DB 148 CACACACATACAGGAGGAGAAATATAAGCTGTAAGAGCTGGGCTGCTTAATGCTGAGCT 207
      TTTTTTTTTTTTTTTTTT
CY 122 CACACACATACAGGAGGAGAAATATAAGCTGTAAGAGCTGCTTAATGCTGAGCT 181
      TTTTTTTTTTTTTTTTTT
DB 208 AACATCATTAATGATTTGGGCTCACTACGAGGCTTTCTGATGGGAGAAACCGTGGTGGC 267
      TTTTTTTTTTTTTTTTTT
CY 182 AACATCATTAATGATTTGGGCTCACTACGAGGCTTTCTGATGGGAGAAACCGTGGTGGC 241
      TTTTTTTTTTTTTTTTTT
DB 268 AAGTGCATTAATGAATCGGCAAGAGAGGAGAGAGGAGGCTTGGGCTATATGGGCTCTTT 327
      TTTTTTTTTTTTTTTTTT
CY 242 AAGTGCATTAATGAATCGGCAAGAGAGGAGAGAGGAGGCTTGGGCTATATGGGCTCTTT 301
      TTTTTTTTTTTTTTTTTT
DB 328 GCTTCTGCTGCTCAATGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
      TTTTTTTTTTTTTTTTTT
CY 302 GCTTCTGCTGCTCAATGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
      TTTTTTTTTTTTTTTTTT

```

```

DB 388 CTCACCTCAAAAGGGGTAATAAGGTTATATATATAAATAGGGATAATATAGAAATATAA 447
      TTTTTTTTTTTTTTTTTT
CY 362 CTCACCTCAAAAGGGGTAATAAGGTTATATATATATATATATATATATATATATATAT 423
      TTTTTTTTTTTTTTTTTT
DB 448 TGTGAGCAAAAGAGGAGAGCAAAAGAGATAGCAAGCTTTAAAAAGAGAGAGATATATAT 497
      TTTTTTTTTTTTTTTTTT
CY 422 TGTGAGCAAAAGAGGAGAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
      TTTTTTTTTTTTTTTTTT
DB 508 TCCATAGAGCTGGGGGGGAGCTGAGAGATATATAAAAAATGAGAGATATAATATATAAG 567
      TTTTTTTTTTTTTTTTTT
CY 482 TCCATAGAGCTGGGGGGGAGCTGAGAGATATAAAAAATGAGAGATATAATATATAAG 542
      TTTTTTTTTTTTTTTTTT
DB 568 GAAACCCGACAGAGATATAAGATATAGAGAGGTTTCTGCTGCTGAGAGAGAGAGAGAG 627
      TTTTTTTTTTTTTTTTTT
CY 542 GAAACCCGACAGAGATATAAGATATAGAGAGAGGTTTCTGCTGCTGAGAGAGAGAGAG 602
      TTTTTTTTTTTTTTTTTT
DB 628 CTCCTGCTGCTGAGAGAGGAGAGAGATATAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
      TTTTTTTTTTTTTTTTTT
CY 602 CTCCTGCTGCTGAGAGAGGAGAGAGATATAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
      TTTTTTTTTTTTTTTTTT
DB 688 TGGCG 692
      TTTT
CY 662 TGGCG 666

```

Search completed: Sun Mar 14 15:24:57 1999

Job time : 7740 secs.

Wed Mar 17 09:43:33 1999

3493 AGCTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGTAACCGCAGGAAGAA 3552
 1154 CATGTGAGCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGTAACCGCAGGAAGAA 1213
 3553 CATGTGAGCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGTAACCGCAGGAAGAA 3612
 1214 TTTCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCGAGGCTCAAAATCAGAGGTTG 1272
 3613 TTTCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCGAGGCTCAAAATCAGAGGTTG 3672
 1274 GCGAAACCCGACGAGCTATAAAGATACCGAGGCTTTCCCTCGAAGCTCCCTCGGTCG 3732
 3673 GCGAAACCCGACGAGCTATAAAGATACCGAGGCTTTCCCTCGAAGCTCCCTCGGTCG 3792
 1334 CTCTCCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAG 1393
 3733 CTCTCCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAG 3792
 1394 CGTGGGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTGCTTCGCTC 1453
 3793 CGTGGGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTGCTTCGCTC 3852
 1454 CAAGCTGGGCTGTGTGACGAGGAGGCGGCTTACCGGATACCTGTCGCGCTTTCTCCCTTC 1513
 3853 CAAGCTGGGCTGTGTGACGAGGAGGCGGCTTACCGGATACCTGTCGCGCTTTCTCCCTTC 3912
 1514 CTATCCTGTTGAGTCCAAACCGGTAAGACAGAGCTTATCGCCACTGCGCAGCAGCCACTGG 1573
 3913 CTATCCTGTTGAGTCCAAACCGGTAAGACAGAGCTTATCGCCACTGCGCAGCAGCCACTGG 3972
 1574 TAACAGGATTAGCAGAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 1633
 3973 TAACAGGATTAGCAGAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 4032
 1634 TAACAGGATTAGCAGAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 1693
 4033 TAACAGGATTAGCAGAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 4092
 1694 CTTCGGAAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 1753
 4093 CTTCGGAAGGAGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 4152
 1754 TTTTGTGTTTCCAAAGCAGAGATACCGGCTGCTACAGAGTCTTGAAGTGGTGGCC 1813
 4153 TTTTGTGTTTCCAAAGCAGAGATACCGGCTGCTACAGAGTCTTGAAGTGGTGGCC 4212
 1814 GATCTTTTCTAGCGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 1873
 4213 GATCTTTTCTAGCGGCTGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCC 4272
 1874 CATGAGATTATCAAAAGGATCTTCCCTAGATCTTCCCTAGATCTTCCCTAGATCTTCCCT 1933
 4273 CATGAGATTATCAAAAGGATCTTCCCTAGATCTTCCCTAGATCTTCCCTAGATCTTCCCT 4292
 1934 ATCAATCTAAGTATATATGAGTAAACTTGGTCTGACAGTACCAATGCTTAATGAGTGA 1993
 4333 ATCAATCTAAGTATATATGAGTAAACTTGGTCTGACAGTACCAATGCTTAATGAGTGA 4392
 1994 GGCACCTATCTACGAGATCTGCTATTTTCTGCTTCCCTAGATCTTCCCTAGATCTTCCCT 2053
 4393 GGCACCTATCTACGAGATCTGCTATTTTCTGCTTCCCTAGATCTTCCCTAGATCTTCCCT 4452
 2054 GTAGATACTACGATACGGGAGGCTTACCATCTGGCCGAGGCTTACCATCTGGCCGAGGCT 2113
 4453 GTAGATACTACGATACGGGAGGCTTACCATCTGGCCGAGGCTTACCATCTGGCCGAGGCT 4512
 2114 AGACCCAGGCTCAGCGCTCCAGATTTATCAGCAATAAAGCAGCCAGCCAGCCAGCCAGCC 2173
 4513 AGACCCAGGCTCAGCGCTCCAGATTTATCAGCAATAAAGCAGCCAGCCAGCCAGCCAGCC 4572
 2174 GGCAGAAAGTGGTCTGCTGCAACTTTATCGGCTTCCATCCAGTCTATTAATTTGTTGCGGA 2233

2844 CATATTGAAATGTTATTAGAAAAATAAACAATAGGGTTCCGCGCAGATTTCCCGGAAA 2953
 5293 CATATTGAAATGTTATTAGAAAAATAAACAATAGGGTTCCGCGCAGATTTCCCGGAAA 5352
 2954 AATGTCAC 2961
 5353 AATGTCAC 5360

RESULT 2 APBL2KSM 2961 bp DNA circular SYN 10-MAY-1995
 LOCUS pBluescript II KS(-) vector DNA, phagemid excised from lambda
 DEFINITION ZAP11.
 ACCESSION X52329
 NID 958060
 KEYWORDS artificial sequence; cloning vector; expression vector; vector.
 SOURCE cloning vectors.
 ORGANISM Artificial sequences: Cloning vectors.
 1 (bases 1 to 2961)
 Thomas, E.A.
 Direct Submission
 Submitted (20-FEB-1990) to the EMBL/GenBank/DBJ databases. Thomas
 E.A., Stratagene Cloning Systems, 11099 North Torrey Pines Rd., La
 Jolla, CA 92037, USA
 2 (bases 1 to 2961)
 Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.
 Lambda ZAP, a bacteriophage lambda expression vector with in vivo
 excision properties
 Nucleic Acids Res. 16 (15), 7583-7600 (1988)
 8831944
 3 (bases 1 to 2961)
 Altman, M.E., and Short, J.M.
 pBluescript II: gene mapping vectors
 Nucleic Acids Res. 17 (22), 9494 (1989)
 90067967 Location/Qualifiers
 1 2961
 /organism="Cloning vector"
 1 2961
 /note="phagemid pBluescriptII KS(-)"
 BASE COUNT 706 a 758 c 735 g 762 t
 ORIGIN
 Query Match 41.6% Score 2228: DB 32: Length 2961;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 2228: Conservative 0: Mismatches 0: Indels 0: Gaps 0;
 Db 734 GTCGACCTCGAGGGGCGGCGGTACCGAGCTTTTGTCCCTTTAGTGGGGTTAATTGC 793
 QY 3133 GTCGACCTCGAGGGGCGGCGGTACCGAGCTTTTGTCCCTTTAGTGGGGTTAATTGC 3192
 Db 794 GCGCTTGGGTAATCATGGTCTATGTTTCTGTTGAAATTTGTTATCGGCTCACAAT 853
 QY 3193 GCGCTTGGGTAATCATGGTCTATGTTTCTGTTGAAATTTGTTATCGGCTCACAAT 3252
 Db 854 TCACACACATACAGCGCGGAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTACG 913
 QY 3253 TCACACACATACAGCGCGGAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTACG 3312
 Db 914 CTAACTCACAATTAATGGGTTGGGCTACTGCCCGCTTTCCAGTTCGGGAAAGCTGTCGTTG 973
 QY 3313 CTAACTCACAATTAATGGGTTGGGCTACTGCCCGCTTTCCAGTTCGGGAAAGCTGTCGTTG 3372
 Db 974 GCAATGTAATGAAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTGGGGGCTC 1033
 QY 3373 GCAATGTAATGAAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATTGGGGGCTC 3432
 Db 1034 TTGGGCTTGGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
 QY 3433 TTGGGCTTGGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3492
 Db 1094 AGCTCACTCAAAAGCGGTAATAGGGTTATCCACAGATCAGGGGTAACCGCAGGAAGAA 1153

Wed Mar 17 09:43:33 1999

Qy	4873	TAATCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACTCAAC	4932		Cp	5240	GGAAGAGTATGAGTATTAACAATTTCCCGTGTGCGCTTATTCCTTTTTCGGGCATTTT	5181	
Db	9220	CAAGTCAATCTGAGATAGTATGTCGGGACCGAGTTGCTCTTTCGCCGGCGTCAATACG	9279		Db	1453	GCCTTCCTGTTTTTGTCTACCCAGAAACGCTGTGTAAGTAAAGATGCTGAAGATCAGT	1512	
Qy	4933	CAATCTATTTGASAAATAGTATGTCGGGACCGAGTTGCTCTTTCGCCGGCGTCAATACG	4992		Cp	5180	GCCTTCCTGTTTTTGTCTACCCAGAAACGCTGTGTAAGTAAAGATGCTGAAGATCAGT	5201	
Db	9280	CGATATACCGGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTCTTC	9339		Db	1513	TGGGTGCACGAGTGGGTATCATCTCAACAGCGTAAAGATGCTTGAAGATTT	1572	
Qy	4993	CGATATACCGGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTCTTC	5052		Cp	5120	TGGGTGCACGAGTGGGTATCATCTCAACAGCGTAAAGATGCTTGAAGATTT	5061	
Db	9340	GGGCGGAAACTCTCAAGAGTCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG	9399		Db	1573	TTCCGCGGAGAGACGTTTTCCAAATGATGAGCACITTTAAAGTCTGTGATGCTGTGCGG	1632	
Qy	5053	GGGCGGAAACTCTCAAGAGTCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG	5112		Cp	5060	TTCCGCGGAGAGACGTTTTCCAAATGATGAGCACITTTAAAGTCTGTGATGCTGTGCGG	5001	
Db	9400	TGAATCAAGTCTTCAAGAGTCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG	9459		Db	1633	TATTAATCCGCTATTGAGCGCGGCGGAGAGCACTCGTGTGAGATGCTGTGAGATTT	1692	
Qy	5113	TGAATCAAGTCTTCAAGAGTCTTACCGTGTGTGAGATCCAGTTCGATGAACCCACTCG	5172		Cp	5000	TATTAATCCGCTATTGAGCGCGGCGGAGAGCACTCGTGTGAGATGCTGTGAGATTT	4941	
Db	9460	AGGAAGGCAAAATGCGGCAAAAGGGAATTAAGGCGGACAGGAAATGTTGAATACTCAT	9519		Db	1693	ATGACTTGGTGTGAGTACTCACCAGTACAGAAAGAGCATTTACGGAATGCGATGACATTA	1752	
Qy	5173	AGGAAGGCAAAATGCGGCAAAAGGGAATTAAGGCGGACAGGAAATGTTGAATACTCAT	5232		Cp	4940	ATGACTTGGTGTGAGTACTCACCAGTACAGAAAGAGCATTTACGGAATGCGATGACATTA	4881	
Db	9520	ACTCTCTCTTTTCAATATTTAAGGCAATTTATCAGGTTTATGCTCATGAGCGGATA	9579		Db	1753	GAGAATTTATGAGTGTCTCCATATACCATGAGTGTATACACTGCGGCGCACTTACTTGA	1812	
Qy	5233	ACTCTCTCTTTTCAATATTTAAGGCAATTTATCAGGTTTATGCTCATGAGCGGATA	5292		Cp	4880	GAGAATTTATGAGTGTCTCCATATACCATGAGTGTATACACTGCGGCGCACTTACTTGA	4821	
Db	9580	CATATTTGAATGTTTGAATAAATAAAGGGAATTAAGGCGGACAGTATTTCCCGGAAA	9639		Db	1813	CAAGGATCGGAGGCGGAGGAGCTTAACCGCTTTTTTTCACAAACATGGGGATCATGTAA	1872	
Qy	5293	CATATTTGAATGTTTGAATAAATAAAGGGAATTAAGGCGGACAGTATTTCCCGGAAA	5352		Cp	4820	CAAGGATCGGAGGCGGAGGAGCTTAACCGCTTTTTTTCACAAACATGGGGATCATGTAA	4761	
Db	9640	AGTGGCAC 9647			Db	1873	CTCGCTTCTGATCGTTGGGAAACCGGAGCTTAAGTGAAGCCCTACCAACAGCAGCGTGA	1932	
Qy	5353	AGTGGCAC 5360			Cp	4760	CTCGCTTCTGATCGTTGGGAAACCGGAGCTTAAGTGAAGCCCTACCAACAGCAGCGTGA	4701	
RESULT	5	AFU25397	9750 bp	DNA	SYN	17-DEC-1997			
LOCUS		AFU25397							
DEFINITION		Expression vector pESP-3, complete sequence.							
ACCESSION		AFU25397							
NID		q2689634							
KEYWORDS		Expression vector pESP-3.							
SOURCE		Expression vector pESP-3.							
ORGANISM		artificial sequence; expression vectors.							
REFERENCE		1. (bases 1 to 9750)							
AUTHORS		Lu, Q.							
TITLE		Direct Submission							
JOURNAL		Submitted (16-SEP-1997) Marketing Analysis, Stratagene, 11011 North							
COMMENT		Torrey Pines Road, La Jolla, CA 92037, USA							
FEATURES		On Dec 17, 1997 this sequence version replaced gi:2623066							
		Location/Qualifiers							
		1..9750							
		/organism="Expression vector pESP-3"							
		/specific_host="Schizosaccharomyces pombe"							
		/db_xref="taxon:57775"							
		/plasmid="pESP-3"							
BASE COUNT		2919 a 1959 c 1983 g 2889 t							
ORIGIN									
		Query Match 41.6%; Score 2228; DB 32; Length 9750;							
		Best Local Similarity 100.0%; Pred No. 0.00e+00;							
		Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Db	1273	GTGGCACTTTTGGGGAATGTGCGGAAACCGCTTATTTGTTTATTTTCTTAATACATT	1332		Db	2293	AGATTGATTTAAACCTTCATTTTAAATTTAAAGAGTCTAGGAGATCTTTTCTA	2352	
Cp	5360	GTGGCACTTTTGGGGAATGTGCGGAAACCGCTTATTTGTTTATTTTCTTAATACATT	5301		Cp	4340	AGATTGATTTAAACCTTCATTTTAAATTTAAAGAGTCTAGGAGATCTTTTGTATA	4281	
Db	1333	CAATATGATTCGGCTCATGACACATACCTGTAAATGCTTCAATATATGAAAA	1392		Db	2353	ATCTATGACCAAAATCCCTTAACGTGAGTTTCTGTCACAGGAGGAGGAGGAGGAG	2412	
Cp	5300	CAATATGATTCGGCTCATGACACATACCTGTAAATGCTTCAATATATGAAAA	5241		Cp	4280	ATCTATGACCAAAATCCCTTAACGTGAGTTTCTGTCACAGGAGGAGGAGGAGGAG	4221	
Db	1393	GGAAGAGTATGAGTATTAACAATTTCCGTTGCGGCTTATTCCTTTTTCGGGCATTTT	1452		Db	2413	AAAAGATCAAGAGTCTTCTTGAGATCCCTTTTCTCGGGGTATCTGCTGCTGCAAA	2472	
					Cp	4220	AAAAGATCAAGAGTCTTCTTGAGATCCCTTTTCTCGGGGTATCTGCTGCTGCAAA	4161	
					Db	2473	CAAAAAACCAACCCCTACAGCGGTTGTTTCTCGGGATCAAGAGTACCAATCTTTT	2532	

US-09-020-716-3.rge

Wed Mar 17 09:43:33 1999

Cp	5180	GCCTTCCTGTTTTCCTCACCACAAACGCTGGTGAAGATGCTGAAGATCAGT	5121	Db	2593	CGTAGTAGGCGACACCTTCAAGAACTCTAGACACCGCTACATACCTGCTGCTAA	2652					
Db	1513	TGGGTGCACAGATGGTTCATACGAACTGGATCTACACGCGTAACTCTTGAGAGTT	1572	Cp	4040	CGTAGTAGGCGACACCTTCAAGAACTCTAGACACCGCTACATACCTGCTGCTAA	3981					
Cp	5120	TGGGTGCACAGATGGTTCATACGAACTGGATCTACACGCGTAACTCTTGAGAGTT	5061	Db	2653	TCTGTTACAGTGGCTGCTGCGATGAAGTGGTGTCTTACCGGTTGAGTCAAA	2712					
Db	1573	TTCGCCCGCAAGAGTTTTCATATGATGACACTTTTAAAGTCTGCTATGTGGCGGG	1632	Cp	3980	TCTGTTACAGTGGCTGCTGCGATGAAGTGGTGTCTTACCGGTTGAGTCAAA	3921					
Cp	5060	TTCGCCCGCAAGAGTTTTCATATGATGACACTTTTAAAGTCTGCTATGTGGCGGG	5001	Db	2713	GAGGATAGTTACCGGATAAGGCGCAGCGTGGGCTGAACGGGGTTCGTCACACAGC	2772					
Db	1533	TATTAACCGGATAAGACGCGGGAAGAGCACTGCTGCGCGATACACTATTTCTCAGA	1692	Cp	3920	GACGATAGTTACCGGATAAGGCGCAGCGTGGGCTGAACGGGGTTCGTCACACAGC	3961					
Cp	5000	TATTAACCGGATAAGACGCGGGAAGAGCACTGCTGCGCGATACACTATTTCTCAGA	4941	Db	2773	CCAGCTTGGAGCGAAGCACTACACGGAAGTACAGCTACAGCGTGAAGTATGAGAA	2832					
Db	1693	ATGACTTGGTGTAGTACTCACCAGTACAGAAAGCACTTTACGGATGGCATGACAGTAA	1752	Cp	3860	CCAGCTTGGAGCGAAGCACTACACGGAAGTACAGCTACAGCGTGAAGTATGAGAA	3801					
Cp	4940	ATGACTTGGTGTAGTACTCACCAGTACAGAAAGCACTTTACGGATGGCATGACAGTAA	4881	Db	2833	CGCCACGCTTCCGGAAGGAGGAGGAGGAGGAGTATCCGGTAAGCGGAGGCTCGGAA	2892					
Db	1753	GAGAAATATGAGTGGTGTAGTACTCACCAGTACAGAAAGCACTTTACGGATGGCATGACAGTAA	1812	Cp	3800	CGCCACGCTTCCGGAAGGAGGAGGAGGAGGAGTATCCGGTAAGCGGAGGCTCGGAA	3741					
Cp	4880	GAGAAATATGAGTGGTGTAGTACTCACCAGTACAGAAAGCACTTTACGGATGGCATGACAGTAA	4821	Db	2893	CAGGAGAGCGACGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTAAGTCTGCTG	2952					
Db	1813	CAAGGATGGGAGGACCAAGGAGCTAACCGCTTTTTCGACAAATATGGGGATCATGTAA	1872	Cp	3740	CAGGAGAGCGACGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTAAGTCTGCTG	3681					
Cp	4820	CANCAATGAGGAGGACCAAGGAGCTAACCGCTTTTTCGACAAATATGGGGATCATGTAA	4761	Db	2953	GCTTTCCGACCTCTGACTTGGAGGTCGATTTTGTGATGCTGCTCAGGGGCGGAGCC	3012					
Db	1873	CTCGCTTGTGTTGGGAAACCGGAGCTGAATGAAGCCATACCAAGACGAGCGGTGACA	1932	Cp	3680	GCTTTCCGACCTCTGACTTGGAGGTCGATTTTGTGATGCTGCTCAGGGGCGGAGCC	3621					
Cp	4760	CTCGCTTGTGTTGGGAAACCGGAGCTGAATGAAGCCATACCAAGACGAGCGGTGACA	4701	Db	3013	TATGNAAGCCACGAGCAACCGGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTTTG	3072					
Db	1933	CAAGGATGGGAGGACCAAGGAGCTAACCGCTTTTTCGACAAATATGGGGATCATGTAA	1992	Cp	3620	TATGNAAGCCACGAGCAACCGGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTTTG	3561					
Cp	4700	CAAGGATGGGAGGACCAAGGAGCTAACCGCTTTTTCGACAAATATGGGGATCATGTAA	4641	Db	3073	CTCACAATGTTTCTTCTGCTGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTTTG	3132					
Db	1993	CTCAGCTTCCCGGCAACAAATATAGACTCGATGGAGCGGATGAAAGTTTCAGGACCA	2052	Cp	3560	CTCACAATGTTTCTTCTGCTGCTTTTACGGTTCCTGGCTTTTGTGCTGCTTTTG	3501					
Cp	4640	CTCAGCTTCCCGGCAACAAATATAGACTCGATGGAGCGGATGAAAGTTTCAGGACCA	4581	Db	3133	AGTCACTGATATACCGTTCGCGCAGCGCAAGCGAGCGAGCGAGCTGAGCGAGG	3192					
Db	2053	TTCGCGCTCGGCGCTTCCGCGCTGCTGCTTTTATGCTGATTAATCTGAGCGGTGAGC	2112	Cp	3500	AGTCACTGATATACCGTTCGCGCAGCGCAAGCGAGCGAGCGAGCTGAGCGAGG	3441					
Cp	4580	TTCGCGCTCGGCGCTTCCGCGCTGCTGCTTTTATGCTGATTAATCTGAGCGGTGAGC	4521	Db	3193	AAGCGGAAGAGCGCGCAATACGCAACCGCTCTCCCGCGCGTTGGCGGATTCATTAAT	3252					
Db	2113	GTGAGGCTCGGCTGATCATTTGAGAGTATGAGGCTGAGTGGTAAAGCTTCGCTATCGTAG	2172	Cp	3440	AAGCGGAAGAGCGCGCAATACGCAACCGCTCTCCCGCGCGTTGGCGGATTCATTAAT	3361					
Cp	4520	GTGAGGCTCGGCTGATCATTTGAGAGTATGAGGCTGAGTGGTAAAGCTTCGCTATCGTAG	4461	Db	3253	CAGCTGCGACGAGCAGGTTTCCCGACTGGAAAGCGGAGCGGAGTGAAGCAATTAATG	3312					
Db	2173	TATCTACAGACGCGGAGTCAAGCAACTATGATGAAGCAATAGACAGATCGCTGAGA	2232	Cp	3380	GCAGTGCAGCAGCAGGTTTCCCGACTGGAAAGCGGCGAGTGAAGCAATTAATG	3321					
Cp	4460	TATCTACAGACGCGGAGTCAAGCAACTATGATGAAGCAATAGACAGATCGCTGAGA	4401	Db	3313	TCAGTTAGCTCACTCATTAGCACCGGCTTTACACTTTATGCTTCGCGCTGCTATGT	3372					
Db	2233	TAGTGGCTTCACTGATTAAGCAATTTGGTAACTGTCAGACCAAGTTTACTCATATACTTT	2292	Cp	3320	TCAGTTAGCTCACTCATTAGCACCGGCTTTACACTTTATGCTTCGCGCTGCTATGT	3261					
Cp	4400	TAGTGGCTTCACTGATTAAGCAATTTGGTAACTGTCAGACCAAGTTTACTCATATACTTT	4341	Db	3373	TGTTGGGAATTTGAGCGGATAACAATTTACAGCGGGAACAGCTATGAGCAATTAATG	3432					
Db	2293	AGATTGATTTAAAGTCACTTTTAAATTTAAAGGATCTAGTGAAGTCTCTTTTGATA	2352	Cp	3260	TGTTGGGAATTTGAGCGGATAACAATTTACAGCGGGAACAGCTATGAGCAATTAATG	3201					
Cp	4340	AGATTGATTTAAAGTCACTTTTAAATTTAAAGGATCTAGTGAAGTCTCTTTTGATA	4281	Db	3433	CCAAGCGGCAATTAAGCTTCACTTAAGGGAACAAAGCTGGGTAAAGCGGCGGCTCG	3492					
Db	2353	ATCTAAGACAAATCCCTTAAGCTGAGTTTTCGTTCCACTGAGCGTCAAGCCCGTAG	2412	Cp	3200	CCAAGCGGCAATTAAGCTTCACTTAAGGGAACAAAGCTGGGTAAAGCGGCGGCTCG	3141					
Cp	4280	ATCTAAGACAAATCCCTTAAGCTGAGTTTTCGTTCCACTGAGCGTCAAGCCCGTAG	4221	Db	3493	AGGTGCGAC	3500					
Db	2413	AAAAGATCAAGAGTCTTCTTGAGATCCCTTTTTCGCGGTAACTGCTGCTTGCAGAA	2472	Cp	3140	AGGTGCGAC	3133					
Cp	4220	AAAAGATCAAGAGTCTTCTTGAGATCCCTTTTTCGCGGTAACTGCTGCTTGCAGAA	4161	RESULT	8	EVU67875	10765 bp	DNA	circular	SYN	16-DEC-1997	
Db	2473	CAAAAACACCGCTTACAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTT	2532	LOCUS								
Cp	4160	CAAAAACACCGCTTACAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTT	4101	DEFINITION								
Db	2533	ITCGGAAGTAACTGGCTTTCAGCAGAGCGCAGATACCAAACTGCTGCTTCTAGTACG	2592	ACCESSION								
Cp	4100	ITCGGAAGTAACTGGCTTTCAGCAGAGCGCAGATACCAAACTGCTGCTTCTAGTACG	4041	NID								
				KEYWORDS								
				SOURCE								

Wed Mar 17 09:43:33 1999

Cp 4100 CAAAAAACACAGCGCTACGAGCGGTGGTTGTTTGGCGGATCAAGAGCTACCAACTCTTT 4101
 Db 2573 TTCCGACGCTAAGTGGCTTCAGCAGAGCGCAGATACCAAACTACTGCTTCTAGTGTAGC 2592
 Cp 4100 TTCCGACGCTAAGTGGCTTCAGCAGAGCGCAGATACCAAACTACTGCTTCTAGTGTAGC 4041
 Db 2593 CTAAGTGTAGTGTACCTTCAAGAACCTGTGTAGCAGCGGCTACATACCTCGCTGCTAA 2652
 Cp 4040 CGTAGTGTAGGCGACCACTTCAAGAACCTGTGTAGCAGCGGCTACATACCTCGCTGCTAA 3981
 Db 2653 TCTGTTTACGAGTGGCTGCTGCGAGTGGCGATAGTGTGTGTACCGGTTGGACTCAA 2712
 Cp 3981 TCTGTTTACGAGTGGCTGCTGCGAGTGGCGATAGTGTGTGTACCGGTTGGACTCAA 3921
 Db 2712 GACGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 2772
 Cp 3921 GACGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3861
 Db 2773 CGAGCTTGGAGGAGCGAGCTACACCGAAGTGTAGATACCTACAGCGTGTAGCTATGAGAA 2832
 Cp 3861 CGAGCTTGGAGGAGCGAGCTACACCGAAGTGTAGATACCTACAGCGTGTAGCTATGAGAA 3801
 Db 2832 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 2892
 Cp 3801 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3741
 Db 2892 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 2952
 Cp 3741 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3681
 Db 2952 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3012
 Cp 3681 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3621
 Db 3012 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3072
 Cp 3621 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3561
 Db 3072 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3132
 Cp 3561 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3501
 Db 3132 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3192
 Cp 3501 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3441
 Db 3192 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3252
 Cp 3441 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3381
 Db 3252 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3312
 Cp 3381 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3321
 Db 3312 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3372
 Cp 3321 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3261
 Db 3372 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3432
 Cp 3261 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3201
 Db 3432 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3492
 Cp 3201 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3141
 Db 3492 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3500
 Cp 3141 GATGATAGTGTACGAGTAAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGCTGCACACAGC 3133

RESULT 9 SYNBLKSPV 2964 bp DNA circular SYN 04-JUN-1993
 LOCUS

DEFINITION BlueScribe KS Plus cloning vector.
 ACCESSION L08785
 NID g310729
 KEYWORDS Synthetic construct DNA.
 ORGANISM artificial construct.
 SOURCE 1 (bases 1 to 2964)
 REFERENCE Gilbert, W.
 TITLE Obtained from VecBase 3.0
 JOURNAL Unpublished (1991)
 COMMENT These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program.
 ENTRY BLUEKSP
 TITLE BlueScribe KS Plus - Cloning vector
 DATE 28-JAN-1987
 #sequence 02-FEB-1987
 #sequence 04-MAR-1987
 #sequence 03-APR-1987
 ACCESSION V00078
 SOURCE artificial
 REFERENCE
 #number 1
 #authors Fernandez J.M., Short J.M., Kenschaw M., Hare W., Surde J
 #journal Gene (1987) in press
 #citation Sequence data from Stratagene
 #comment sequence correction according to Stratagene COMMENT
 obtained from Stratagene on floppy disc.
 Revised 02-FEB-1987 by F. Pfeiffer;
 1409/10 'AT' to 'TA' to match revised sequence of PAR322
 Revised 4-MAR-1987 to match sequence of pUC19 on request
 of Stratagene
 Polylinker region revised 03-APR-1987 according to Stratagene
 COMMENT
 The stand shown corresponds to pUC19c.
 As in the published sequence of pUC19c, The M13mp18 lacZ region
 is on the complementary strand.
 COMMENT
 This vector contains the fl origin so that the plus strand
 can be obtained upon fl superinfection.
 KEYWORDS
 CROSSREFERENCE
 #parent
 Vecbase(3):BlueM13p
 #parent
 Vecbase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3.
 VecSource(3):bGalks, GenBank(50):PFI
 #brother
 Vecbase(3):BlueKSm, Vecbase(3):BlueSkp
 PARENT
 Features of BlueKsp (2964 bp)
 residue source
 3- 458 5943-5488 (c) phage fl
 460- 624 236- 400 pUC19c
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKs-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-2964 448-2617 pUC19c
 Conflict (cfl) and Mutations (mut): none
 PARENT
 Features of BlueKsp (2964 bp)
 residue source
 3- 458 5943-5488 (c) phage fl
 460- 624 236- 400 pUC19c
 626- 645 1- 20 T7 promoter
 653- 760 108- 1 (c) BlueKs-polylinker
 772- 791 20- 1 (c) T3 promoter
 795-1031 237- 1 (c) pUC19
 1032-2964 2586- 754 (c) pUC19
 Conflict (cfl) and Mutations (mut): none
 FEATURE

Wed Mar 17 09:43:33 1999

```

QY 4993 GGATAATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATCTTGGAAACGTTCTTC 5052
DB 2654 GGGGCGAAACCTTCAGAGATCTTACGCTGTGTGAGATCACTTCGATGTAACCCACTCG 2713
QY 5053 GGGGCGAAACCTTCAGAGATCTTACGCTGTGTGAGATCACTTCGATGTAACCCACTCG 5112
DB 2714 TGCACCAACTGATCTTCAGCATCTTTTACTTTCACCAAGGTTTCGGTGAGCAAAAC 2773
QY 5113 TGCACCAACTGATCTTCAGCATCTTTTACTTTCACCAAGGTTTCGGTGAGCAAAAC 5172
DB 4774 AGAAGGCAAAATCGCCCAAAAGGAATAAGGGGACACGGAATGTTGAATACTCAT 2833
QY 5173 AGAAGGCAAAATCGCCCAAAAGGAATAAGGGGACACGGAATGTTGAATACTCAT 5232
DB 2834 ACTCTTCCTTTTCAATATTATTAAGCATTTATCAGGGTTATGTCTCATGAGCGGATA 2893
QY 5233 ACTCTTCCTTTTCAATATTATTAAGCATTTATCAGGGTTATGTCTCATGAGCGGATA 5292
DB 2894 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA 2953
QY 5293 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA 5352
DB 2954 ATGGCCAC 2961
QY 5353 AGTGGCAC 5360

```

```

RESULT 10 SYNBLKSW 2964 bp DNA circular SYN 04-JUN-1993
LOCUS BlueScribe KS Minus cloning vector.
DEFINITION L08784
ACCESSION g310728
KEYWORDS Synthetic construct DNA.
SOURCE Synthetic construct.
ORGANISM Artificial sequence.
REFERENCE 1 (bases 1 to 2964)
AUTHORS Gilbert W.
TITLE Obtained from VecBase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program.
BlueScribe KS Minus - Cloning vector
ENTRY BlueScribe KS Minus - Cloning vector
TITLE BlueScribe KS Minus - Cloning vector
DATE 29-JAN-1987
#sequence 02-FEB-1987
#sequence 04-MAR-1987
#sequence 03-APR-1987
ACCESSION V00077
SOURCE artificial
REFERENCE
#number 1
#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorge J.
#journal Gene (1987) in press
#citation Sequence data from StrataGene
#comment sequence correction according to StrataGene COMMENT
Obtained from StrataGene on floppy disc.
Revised 02-FEB-1987 by F. Pfeiffer.
1409/10 'At' to 'TA' to match revised sequence of pPC12
Revised 4-MAR-1987 to match sequence of pUC19 on request
of StrataGene
Polylinker region revised 03-APR-1987 according to StrataGene
COMMENT
The stand shown corresponds to pUC19c.
As in the published sequence of pUC19c, The M13mpl9 lacZ region
is on the complementary strand.
COMMENT
This vector contains the fl origin so that the minus strand
can be obtained upon fl superinfection.
KEYWORDS
CROSSREFERENCE

```

```

#parent
VecBase(3):BlueM13m
#parent
VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,
VecSource(3):bGalKS, GenBank(50):PF1
#brother
VecBase(3):BlueKSp, VecBase(3):BlueKSm
PARENT
Features of BlueKSm (2964 bp)
residue source
3-458 5488-5943 phage fl
460-624 236-490 pUC19c
626-645 1-20 T7 promoter
653-760 108-1 (c) BlueKS-polylinker
772-791 20-1 (c) T3 promoter
795-2964 448-2617 pUC19c
Conflict (cfl) and Mutations (mut): none
PARENT
Features of BlueKSm (2964 bp)
residue source
3-458 5488-5943 phage fl
460-624 449-285 (c) pUC19
626-645 1-20 T7 promoter
653-760 108-1 (c) BlueKS-polylinker
772-791 20-1 (c) T3 promoter
795-1031 237-1 (c) pUC19
1032-2964 2686-754 (c) pUC19
Conflict (cfl) and Mutations (mut): none
FEATURE 643 start of T7-RNA synthesis
774 (c) start of T3-RNA synthesis
1976-2764 789-1 (c) Ap-R; b-lactamase
POLYLINKER
KpnI-draII-ApaI-XhoI-SalI-HindIII-EcoRV-EcoFI-PstI-
SmaI-BamHI-SpeI-XbaI-NciI-XmaII-BstXI-SacII-SacI SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY BlueKSm #length 2964 #checksum 1589.
FEATURES
source
1. 2964
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 708 a 756 c 735 g 765 t
ORIGIN
Query Match 41.5%, Score 2224, DB 32; Length 2964:
Best Local Similarity 99.9%; Pred No. 0.00e+00;
Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0:
DB 734 GTCGACCTCGAGGGGGCGGGTACCCAGCTTTTGTTCCTTTAGTGGGTTATTC 793
QY 3133 GTCGACCTCGAGGGGGCGGGTACCCAGCTTTTGTTCCTTTAGTGGGTTATTC 3192
DB 794 GAGCTTGGCGTAATCATGTCTAGCTTTTCCTGTGTGAATTTTATCGGTCAAT 853
QY 3193 GCGCTTGGCGTAATCATGTCTAGCTTTTCCTGTGTGAATTTTATCGGTCAAT 3252
DB 854 TCCACACACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGTGCTATGATGAG 913
QY 3253 TCCACACACATACGAGCGGGAAGCATAAAGTGTAAAGCTTGGGTGCTATGAG 3312
DB 914 CTAACTACATTAATTTGCGTTCGCTACATCTTCGCTGCTGCTGCTGCTGCTG 973
QY 3313 CTAACTACATTAATTTGCGTTCGCTACATCTTCGCTGCTGCTGCTGCTGCTG 3372
DB 974 CCAGCTGCATTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
QY 3373 CCAGCTGCATTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3432
DB 1034 TTCCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
QY 3433 TTCCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3492

```


Wed Mar 17 09:43:33 1999

1020 East Meadow Circle, Palo Alto, CA 94303, USA
This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94304, USA. To place an order call (415) 424-8222 or (800) 552-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 552-2566, extension 3.
This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH; this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 552-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES		source		Location/Qualifiers			
		1..2967		/organism="Cloning vector pDIRECT"			
		/db_xref="taxon:31799"					
BASE COUNT		726 a	756 c	735 g	750 t		
ORIGIN							
Query Match		41.5%		Score 2224; DB 32; Length 2967;			
Best Local Similarity		99.9%;		Pred. No. 0.00e+00;			
Matches 2256;		Conservative		0; Mismatches 2;		Indels 0; Gaps 0;	
Db	700	CTGGCACTTTTGGGCAATGTGGGCAATGCTGATATTTGTTTCTAAATACATT	759				
Cp	5460	GTGGCACTTTTGGGCAATGTGGGCAATGCTGATATTTGTTTCTAAATACATT	5301				
Db	760	CAATATGATGCGCTCATGACACATACCTGATATATGCTTCAATATATTAATAA	819				
Cp	5300	CAATATGATGCGCTCATGACACATACCTGATATATGCTTCAATATATTAATAA	5241				
Db	820	GGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	879				
Cp	5440	GGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	5181				
Db	880	GCTTTCCTGTTTGTGTCACCCAGAACGCTGCTGAAATGATGATGATGATGATG	939				
Cp	5180	GCTTTCCTGTTTGTGTCACCCAGAACGCTGCTGAAATGATGATGATGATGATG	5121				
Db	940	TGGTGGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	999				
Cp	5120	TGGTGGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	5061				
Db	1000	TGGTGGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	1059				
Cp	5060	TGGTGGACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG	5001				
Db	1060	TAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1119				
Cp	5000	TAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4941				
Db	1120	ATGACTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1179				
Cp	4940	ATGACTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4881				
Db	1180	GAGAAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1239				
Cp	4880	GAGAAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	4821				
Db	1240	CAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1299				
Cp	4820	CAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4761				
Db	1300	CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1359				
Cp	4760	CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4701				
Db	1360	CCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1419				
Cp	4700	CCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4641				

Db	1420	CTCTAGCTTCCCGGCAACAATAATAGATGATGATGATGATGATGATGATGATGATG	1479		
Cp	4640	CTCTAGCTTCCCGGCAACAATAATAGATGATGATGATGATGATGATGATGATGATG	4581		
Db	1480	TTCGCTGCTGCGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1539		
Cp	4580	TTCGCTGCTGCGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4521		
Db	1540	GTGGTCTCGGCTGATCATTCAGCAGCTGCGGCGAGATGATGATGATGATGATG	1599		
Cp	4520	GTGGTCTCGGCTGATCATTCAGCAGCTGCGGCGAGATGATGATGATGATGATG	4461		
Db	1600	TATCTACACGCGGGAGTCAGCAACTATGATGAACGAAATAGACAGATGCTGAGA	1659		
Cp	4460	TATCTACACGCGGGAGTCAGCAACTATGATGAACGAAATAGACAGATGCTGAGA	4401		
Db	1660	TAGTGCTCTCACTGATTAAAGCATTTAGTTCAGCAGCAAGTTTACTCATATATCT	1719		
Cp	4400	TAGTGCTCTCACTGATTAAAGCATTTAGTTCAGCAGCAAGTTTACTCATATATCT	4341		
Db	1720	AGATTGATTTAAAGCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGA	1779		
Cp	4340	AGATTGATTTAAAGCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTTGA	4281		
Db	1780	ATCTCATGACCAAAATCCCTTAACTGATGATTTTCTGCGCTGATGATGATGATG	1839		
Cp	4280	ATCTCATGACCAAAATCCCTTAACTGATGATTTTCTGCGCTGATGATGATGATG	4221		
Db	1840	AAAGATCAAGATGCTTCTGATGATCCTTTTCTGCGCTGATGATGATGATGATG	1899		
Cp	4220	AAAGATCAAGATGCTTCTGATGATCCTTTTCTGCGCTGATGATGATGATGATG	4161		
Db	1900	CAAAAACCAACGCTACCAAGCGGTGCTTTGTTTGGGATCAAGAGCTACCACTCT	1959		
Cp	4160	CAAAAACCAACGCTACCAAGCGGTGCTTTGTTTGGGATCAAGAGCTACCACTCT	4101		
Db	1960	TTCGAGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG	2019		
Cp	4100	TTCGAGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG	4041		
Db	2020	CGTAGTTAGGCGCACCTTCAAGAACTCTGATGATGATGATGATGATGATGATG	2079		
Cp	4040	CGTAGTTAGGCGCACCTTCAAGAACTCTGATGATGATGATGATGATGATGATG	3981		
Db	2080	TCTGTTTACAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATG	2139		
Cp	3980	TCTGTTTACAGTGGCTGCTGATGATGATGATGATGATGATGATGATGATG	3921		
Db	2140	GAGGATGCTTACCGGATTAAGGCGGCGGTGGGCTGAAAGGCGGCTGATGATG	2199		
Cp	3920	GAGGATGCTTACCGGATTAAGGCGGCGGTGGGCTGAAAGGCGGCTGATGATG	3861		
Db	2200	CCAGCTTGGAGGAGGAGCTACCTGATGATGATGATGATGATGATGATGATG	2259		
Cp	3860	CCAGCTTGGAGGAGGAGCTACCTGATGATGATGATGATGATGATGATGATG	3801		
Db	2260	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2319		
Cp	3800	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3741		
Db	2320	CAGAGAGCGGAGGAGCTTCCAGGCGGAGGAGCTGATGATGATGATGATGATG	2379		
Cp	3740	CAGAGAGCGGAGGAGCTTCCAGGCGGAGGAGCTGATGATGATGATGATGATG	3681		
Db	2380	GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2439		
Cp	3680	GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3621		
Db	2440	TATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2499		
Cp	3620	TATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3561		
Db	2500	CTCAGATGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2559		

Wed Mar 17 09:43:33 1999

Db	1745	CTCAGCTTCTTTCTGCGGTTATCCCTGATCTTGATGAACCGTATTACCCCTTG	2824
Cp	3560	CTCAGCTTCTTTCTGCGGTTATCCCTGATCTTGATGAACCGTATTACCCCTTG	3501
Db	2825	ACTGAGTGTATGATGCTGCGGCTAGCGGAGCGGAGCGGAGTCACTGAGCGAGG	2884
Cp	3500	AGTGAAGTGTATGATGCTGCGGCTAGCGGAGCGGAGCGGAGTCACTGAGCGAGG	3441
Db	2885	AAGCGGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	2944
Cp	3440	AAGCGGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3381
Db	2945	GCAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3004
Cp	3380	GCAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3221
Db	3005	TGAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3064
Cp	3320	TGAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3261
Db	3065	TGAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3124
Cp	3260	TGAGTGTGAGGAGCGGCTCAATACGCAACCGGCTCTCCGCTGCGGCTGCGGCTG	3201
Db	3125	CCAAGCGGCAATTAACCTCACTAAAGGAAACAAAGCTGGGTACCGGCGGCGCTG	3184
Cp	3200	CCAAGCGGCAATTAACCTCACTAAAGGAAACAAAGCTGGGTACCGGCGGCGCTG	3141
Db	3185	AGGTC 3189	
Cp	3140	AGGTC 3136	
RESULT	13		
LOCUS	A58522	7474 bp	DNA circular PAT 05-MAP-1998
DEFINITION	Sequence 26 from Patent WO9638555.		
ACCESSION	A58522		
KEYWORDS	93714138		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 7474)		
AUTHORS	Bogaert, T., Stringham, E. and Vandekerckhove, J.		
TITLE	PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR USE IN THE CONTROL OF CELL BEHAVIOUR		
JOURNAL	Patent: WO 9638555-A 26 05-DEC-1996;		
COMMENT	BOGAERT THIERRY (BE)		
FEATURES	Other publication AU 6123496 961218.		
BASE COUNT	2107 a 1921 c 1655 g 1790 t		
ORIGIN	source		
Query Match	41.5%	Score 2222	DB 22: Length 7474:
Best Local Similarity	100.0%	Pred. No. 0.00e+00:	
Matches	2222	Mismatches 0:	Indels 0: Gaps 0:
Db	5253	CTCAGGCGGCGGCTGAGTACCGAGCTTTTGTTCCTTTAGTGAGGGTTAATTCGCGCTT	5312
Qy	3139	CTCAGGCGGCGGCTGAGTACCGAGCTTTTGTTCCTTTAGTGAGGGTTAATTCGCGCTT	3198
Db	5313	GGGTATCATGTGATAGCTGTTTCCTGTTGAAATTTGTTATCCGCTCAACATTCACA	5372
Qy	3199	GGGTATCATGTGATAGCTGTTTCCTGTTGAAATTTGTTATCCGCTCAACATTCACA	3258
Db	5373	CAACATACGAGCGGAGGATTAAGTGTAAAGCTTGGGCTGCTTAAGTGTAAAGCT	5432
Qy	3259	CAACATACGAGCGGAGGATTAAGTGTAAAGCTTGGGCTGCTTAAGTGTAAAGCT	3318

Wed Mar 17 09:43:33 1999

```

KEYWORDS      Cloning vector pFUN
SOURCE         Cloning vector pFUN
ORGANISM       artificial sequence; cloning vectors.
REFERENCE      1 (bases 1 to 8072)
AUTHORS        Poquet, I., Ehrlich, S.D. and Gruss, A.
TITLE          An export-specific reporter designed for gram-positive bacteria:
               application to Lactococcus lactis
JOURNAL        J. Bacteriol. 180 (7), 1904-1912 (1998)
MEDLINE        98196737
REFERENCE      2 (bases 1 to 8072)
AUTHORS        Poquet, I. and Gruss, A.
TITLE          Direct Submission
JOURNAL        Submitted (15-DEC-1997) Laboratoire de Genetique Appliquee-URLGA,
               Institut National de la Recherche Agronomique, CRJ, Jouy en Josas
               78352, France
FEATURES       Location/Qualifiers
               1..8072
               /organism="Cloning vector pFUN"
               /db_xref="taxon:74538"
               /note="plasmid vector that creates translational fusions
               with a nucleic acid reporter without a signal peptide"
misc_feature   1..4580
               /note="pL252 moiety: similar to sequence deposited under
               GenBank Accession Number AF039139"
misc_feature   4581..7472
               /note="similar to BlueScript"
terminator     7473..7507
               /note="trpA terminator"
misc_feature   7512..7547
               /note="multiple cloning site"
gene           <7543..8010
               /note="truncated nuc gene potentially encoding deltaSP-Nuc
               ORF"
CDS            /gene="delta-nuc"
               <7543..8010
               /gene="delta-nuc"
               /note="nuclease from Staphylococcus aureus devoid of its
               signal peptide and the thirteen amino acids at the
               N-terminus of mature protein"
               /codon_start=1
               /transl_table=1
               /product="deltaSP-Nuc"
               /db_xref="PIB:G3043925"
               /translation="DPTVYSATSKLHKPEATLIIKIDGTVKLMYKQPMTERLLL
               VDPETKPKKGVKYGPEASATKKWENAKKIEVEDKQRTDKYGRGLAYVADG
               KWNVEALVPGGLAKVAYVYKPNNTHEOHLKSEAKKEKLINSEDNADSSQ"
BASE COUNT    2717 a 1474 c 1694 g 2187 t
ORIGIN
Query Match      41.5%; Score 22.4; E=32; Length 8072;
Best Local Similarity 44.0%; Pred No. 0.00e+00;
Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
Db 5240 GTGGCACTTTTCGGGAAATGTGGCGGAACCCCTATTGTTTATTTCTAATACATT 5299
Cp 5360 GTGGCACTTTTCGGGAAATGTGGCGGAACCCCTATTGTTTATTTCTAATACATT 5301
Db 5360 CAATATGATCGCTCATGAGACATAATACCGCTGATAATGCTCAATAATATTGAAAA 5359
Cp 5360 CAATATGATCGCTCATGAGACATAATACCGCTGATAATGCTCAATAATATTGAAAA 5241
Db 5360 GGAAGAGTATGAGTATTAACATTTCCGTGGCCCTATTCCCTTTTTCGGGCATTTT 5419
Cp 5240 GGAAGAGTATGAGTATTAACATTTCCGTGGCCCTATTCCCTTTTTCGGGCATTTT 5181
Db 5420 GCTTTCCTGTTTTCCTGACCGCAACACGCTGGTGAAGAAATAAAGATGCTGAAGATCAGT 5479
Cp 5180 GCTTTCCTGTTTTCCTGACCGCAACACGCTGGTGAAGAAATAAAGATGCTGAAGATCAGT 5121
Db 5480 TGGGTGCAAGATGAGTATTAACATGCTGATGCTGAACAGCGGTGAAGATCCTTTGAGGTT 5539
Cp 5120 TGGGTGCAAGATGAGTATTAACATGCTGATGCTGAACAGCGGTGAAGATCCTTTGAGGTT 5061

```

```

Db 5540 TTGCCCCGGAAGAGCTTTTCCAAATGATGAGCAGCTTTTAAAGTTCTGCTATGTGGCGG 5599
Cp 5060 TTGCCCCGGAAGAGCTTTTCCAAATGATGAGCAGCTTTTAAAGTTCTGCTATGTGGCGG 5001
Db 5600 TATTATCCGTTATGACCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGA 5659
Cp 5000 TATTATCCGTTATGACCGCGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGA 4941
Db 5660 ATGACTTGGTTGAGTACTCACCAGTTCACAGAAAAGCATCTTACGGATGGCATGACAGTAA 5719
Cp 4940 ATGACTTGGTTGAGTACTCACCAGTTCACAGAAAAGCATCTTACGGATGGCATGACAGTAA 4881
Db 5720 GAGAAATATGCACTGCTGCCATAACCATGAGTATACACTTGGCGGCAACTTACTTCTGA 5779
Cp 4880 GAGAAATATGCACTGCTGCCATAACCATGAGTATACACTTGGCGGCAACTTACTTCTGA 4821
Db 5780 CAACGATCGGAGGACCGAAGAGGAGCTAACCGCTTTTTCACAAACATGAGGATCATGATA 5839
Cp 4820 CAACGATCGGAGGACCGAAGAGGAGCTAACCGCTTTTTCACAAACATGAGGATCATGATA 4761
Db 5840 CTGCGCTTGTATGCTTGGGACCGGAGCTGAATGAAGGCAATATCAAAAGCAAGAGCTGACA 5899
Cp 4760 CTGCGCTTGTATGCTTGGGACCGGAGCTGAATGAAGGCAATATCAAAAGCAAGAGCTGACA 4701
Db 5900 CCAGATGCGCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTGGCGGCAACTACTTAA 5959
Cp 4700 CCAGATGCGCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTGGCGGCAACTACTTAA 4641
Db 5960 CTCTAGCTTCCGCGCAACAATTAATAGACTGGATGAGGCGGATAAAGTTGAGAGGACAC 6019
Cp 4640 CTCTAGCTTCCGCGCAACAATTAATAGACTGGATGAGGCGGATAAAGTTGAGAGGACAC 4581
Db 6020 TTCTCGCTCGCCCTTCCGCGCTCGGCTGGCTGTTTATGCTGATATAATATGAGAGGCGG 6079
Cp 4580 TTCTCGCTCGCCCTTCCGCGCTCGGCTGGCTGTTTATGCTGATATAATATGAGAGGCGG 4521
Db 6080 GTGGGTCTCGGGGTATCATTCAGCACTGGGCGGCGAGATGTAAGCCCTCCCGTATGAGTAG 6139
Cp 4520 GTGGGTCTCGGGGTATCATTCAGCACTGGGCGGCGAGATGTAAGCCCTCCCGTATGAGTAG 4461
Db 6140 TTATCTACAGCAGCGGGAGTCAAGCAACTATGGATGAACGAATATAGACAGATCGCTGAGA 6199
Cp 4460 TTATCTACAGCAGCGGGAGTCAAGCAACTATGGATGAACGAATATAGACAGATCGCTGAGA 4401
Db 6200 TAGTCCCTCAGTATTAAAGCATTTGTAAGTGTGAGACCAAGTTTACTCATATATACATTT 6259
Cp 4400 TAGTCCCTCAGTATTAAAGCATTTGTAAGTGTGAGACCAAGTTTACTCATATATACATTT 4341
Db 6260 AGATTGATTAAACCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTGATA 6319
Cp 4340 AGATTGATTAAACCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCTTTTGTGATA 4281
Db 6320 ATCTCATGACCAAAATCCCTTAAGCTGAGTTTCTGTCATGAGAGTGTGAGAGTGTGAGAG 6379
Cp 4280 ATCTCATGACCAAAATCCCTTAAGCTGAGTTTCTGTCATGAGAGTGTGAGAGTGTGAGAG 4221
Db 6380 AAAAGATCAAAAGGATCTTTTGTGAGATCCCTTTTCTGGGGTATCTGCTGCTGCAAAA 6439
Cp 4220 AAAAGATCAAAAGGATCTTTTGTGAGATCCCTTTTCTGGGGTATCTGCTGCTGCAAAA 4161
Db 6440 CAAAAAACCACCGCTACAGCGGTGGTTTGTGTCGGGATCAAGAGTACCAATCTTTT 6499
Cp 4160 CAAAAAACCACCGCTACAGCGGTGGTTTGTGTCGGGATCAAGAGTACCAATCTTTT 4101
Db 6500 TTCCGAAGTAACTGGCTTCAGCAGAGCGAGATACCAATACTGTCCTTCTAGTGTAGC 6559
Cp 4100 TTCCGAAGTAACTGGCTTCAGCAGAGCGAGATACCAATACTGTCCTTCTAGTGTAGC 4041
Db 6560 GGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTGGTGTGTAA 6619
Cp 4040 GGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTGGTGTGTAA 3981

```




Sequence of clone BP2 0.00e+00
Plasmid pWRG3169 enco 0.00e+00
Plasmid pM16-1. 0.00e+00
Plasmid pM16. 0.00e+00
Plasmid pVSV(-) lac 0.00e+00
K.lactis/S. cerevisiae 0.00e+00
Vector p37M1-10D. 0.00e+00
Construct pGEM-HTR(Ka 0.00e+00
Construct pGEM-HTR co 0.00e+00
pDE110. 0.00e+00
Plasmid pMP6-IL2. 0.00e+00
Interleukin-2 expres 0.00e+00
pVE108. 0.00e+00
Plasmid pVE109 used i 0.00e+00
pDE4. plasmid DNA rep 0.00e+00
Plasmid pWRG3196 enco 0.00e+00
Sequence of pVE4. 0.00e+00
Plasmid pCB51 encodin 0.00e+00
Plasmid pCB50 encodin 0.00e+00
Plasmid pCPM1.8 conta 0.00e+00
Human cyclin D1-human 0.00e+00
Shortened c1 gene in 0.00e+00
pDE108. 0.00e+00
pDE102/3 Apal constr 0.00e+00
Plasmid pVE144 used i 0.00e+00
IL-4.Y134D/Y193I fusio 0.00e+00
Plasmid pHCV-157 codi 0.00e+00
pHCV167 sequence. 0.00e+00
Plasmid pHCV-162 codi 0.00e+00
Plasmid pTB72 encodin 0.00e+00
Alphavirus-based euka 0.00e+00

14 2198 41.0 6526 7 040281
15 2198 41.0 7287 39 02042
16 2196 41.0 10930 20 081226
17 2196 41.0 10950 20 081225
18 2195 41.0 14311 24 138744
19 2190 40.9 6824 6 039050
20 2185 40.8 7228 6 050201
21 2166 40.4 4118 35 169189
22 2166 40.4 4118 35 169188
23 2163 40.4 4883 4 025709
24 2163 40.4 4883 8 051192
25 2163 40.4 5585 14 087418
26 2163 40.4 5585 29 161430
27 2163 40.4 5620 4 027489
28 2163 40.4 5620 4 059475
29 2163 40.4 5642 6 051133
30 2167 40.4 6295 39 020419
31 2163 40.4 7050 7 040419
32 2163 40.4 10288 31 171322
33 2163 40.4 13414 31 171321
34 2162 40.3 4454 30 165462
35 2162 40.3 4540 32 169891
36 2162 40.3 4824 20 168975
37 2162 40.3 5399 4 025706
38 2162 40.3 6274 7 043812
39 2162 40.3 6555 9 053874
40 2162 40.3 6926 18 112662
41 2162 40.3 7106 8 047193
42 2162 40.3 7106 15 097424
43 2162 40.3 7298 8 047192
44 2162 40.3 10443 31 171319
45 2162 40.3 16656 21 130787

ALIGNMENTS

RESULT 1
ID T49876 standard; DNA: 5178 BP.
AC T49876;
DE 28-MAR-1997 (first entry)
DE plet-Splice.
KW plet-Splice; plet-trak, tetracycline transactivator; promoter;
KW gene expression; transgenic animal; animal model; drug screening;
KW vector; ds.
OS Synthetic.
FH Key
FT misc_rna
FT Location/Qualifiers
FT complement (2591)
FT /*tag= a
FT /*note= "putative start site of trxn"
FT complement (2616..2622)
FT /*tag= b
FT tata_signal
FT W09640946-Al.
PN 19-DEC-1996.
PD 07-JUN-1996: U10109.
PF 07-JUN-1996: US-474169.
PR (UYVA) UNIV YALE.
PA Schatz DG;
PI WPI: 97-077273/07.
DR Nucleic acid encoding tetracycline transactivator fusion protein -
PT provides rapid and reversible control of gene expression, e.g. for
PT creating animal models for drug screening
PT Example 1: Fig 9b-g, 82pp, English.
PS Vector plet-Splice (T49876) is utilised in the construction of
CC autoregulatory vector Plasmid plet-trak (T49877) Splice-PA was
CC made by ligating the SV40 small T antigen intervening sequence and
CC the SV40 early polyA sequence into plet-Splice. The XhoI-SalI fragment
CC of pUC119-3 (contg. 7 copies of the tet operator upstream of minimal
CC promoter rep) was cloned upstream of the splice/polyA sequence of
CC plet-Splice-PA to form plet-Splice. A modified tetracycline
CC transactivator (ttak) gene was cloned into plet-Splice to form plet-
CC ttak. The construct provides rapid, reversible control of gene
CC expression in eukaryotic cells or transgenic animals, e.g. for
CC heterologous protein, RNA or antisense sequence prodn., or for

M P S R E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on Sun Mar 14 23:15:31 1999. MasPar time 757.38 Seconds
962.844 Million cell updates/sec
Tabular output not generated.

Title: US-09-020-716-3
Description: (1:5360) from US09020716.seq
Perfect Score: 5360
N.A. Sequence: 1 CTAATGTTAAGCGTTAATA... ATTTCGCCGAAAGTGGCAC 5360
N.A. Comp: GAATTAACATTCGCAATTA... TAAAGGCGCTTTACCGGTG

Scoring table:
Gap K
TABLE default
Nmatch STD : Dbase 0: Query 0
Searched: 188442 seqs, 58026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 10.451: Variance 6.768: scale 1.546
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Match	Query Length	ID	Description	Pred. No.
1	2222	41.5	5178	27	T49876	0.00e+00
2	2222	41.5	7474	31	T71320	0.00e+00
3	2220	41.4	6206	27	T49877	0.00e+00
4	2201	41.1	3681	2	Q13578	0.00e+00
5	2202	41.1	4145	7	Q40279	0.00e+00
6	2203	41.1	5314	2	Q13576	0.00e+00
7	2202	41.1	5356	25	T43794	0.00e+00
8	2202	41.1	5534	25	T43137	0.00e+00
9	2198	41.0	3699	39	V14340	0.00e+00
10	2198	41.0	4277	7	Q40280	0.00e+00
11	2195	41.0	4539	14	Q87347	0.00e+00
12	2200	41.0	4792	15	Q84696	0.00e+00
13	2200	41.0	5042	15	Q84694	0.00e+00

SUMMARIES

[illegible]

Db 4965 gacgaccgaattgtcttgcggcggtcgaatacgggataataccggccacacatacagaa 5014
 QV 4969 gacgaccgaattgtcttgcggcggtcgaataacgggataataccggccacacatacagaa 5018
 Db 5015 ctttaaaagtctcatcattggaacacgtctcttcgaggcgaaacactcctaagatcttac 5074
 QV 5019 ctttaaaagtctcatcattggaacacgtctcttcgaggcgaaacactcctaagatcttac 5078
 Db 5075 cctgtatgaatccaggttcgatgaacccactctgcacccaaactgatcttcagatctt 5134
 QV 5079 cctgtatgaatccaggttcgatgaacccactctgcacccaaactgatcttcagatctt 5138
 Db 5135 ttacttccacagcgtttcttggttgagcaaaaacagaggaacaaatgccgcaaaaagg 5194
 QV 5139 ttacttccacagcgtttcttggttgagcaaaaacagaggaacaaatgccgcaaaaagg 5198
 Db 5195 gaataaggcgacacgggaattgtgaatactcactctctctctctctctctctctctctct 5254
 QV 5199 gaataaggcgacacgggaattgtgaatactcactctctctctctctctctctctctctct 5258
 Db 5255 gcaattatcagggtttatctctcatgagcgatacatatttgaaatgatttagaataaata 5314
 QV 5259 gcaattatcagggtttatctctcatgagcgatacatatttgaaatgatttagaataaata 5318
 Db 5315 aacaaatagggttccgcgcacatttcccgcaaaaagtgcac 5356
 QV 5319 aacaaatagggttccgcgcacatttcccgcaaaaagtgcac 5360

RESULT 8
 ID T43137 standard; DNA: 5534 BP.
 AC T43137:
 DT 10-FEB-1997 (first entry)
 DE pUMIGIT sequence including upstream activating sequence.
 KW transgenic plant; gene expression; upstream activating sequence;
 KW UAS; transactivating protein; Gal4; herbicide resistance;
 KW polyhydroxybutyrate; safety; pUMIGIT; ds.
 OS Synthetic.
 PN CA2150039-A.
 PD 09-AUG-1996.
 PE 08-FEB-1995; GB-002455.
 PR (UWMA-) UNIV WARWICK.
 PA Bennett M, May S, Ramsay N:
 PI WPI: 96-486150/49.
 PR Control of genes in transgenic plants - using an upstream activating
 PT sequence activated by a transactivating protein expressed using a
 PI separate promoter
 PS Example: Fig 5F: 48pp: English.
 CC Reporter plasmid pUMIGIT (T43137), or pUAS Minimal promoter Gus
 CC Int Terminator, contains a beta-glucuronidase (GUS) reporter
 CC gene under the control of the 4b5 cauliflower-mosaic virus minimal
 CC promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream
 CC activating sequence (UAS) of pUMIGIT is activatable by yeast
 CC transactivating protein GAL4. In a novel method for controlling
 CC gene expression, a first transgenic plant carrying a gene encoding
 CC a desired phenotype (herbicide resistance of polyhydroxybutyrate
 CC prodn.) operatively linked to a UAS recognition site is pollinated
 CC by a second transgenic plant carrying Gal4 DNA (see also T43136).
 CC The transgene is fully expressed in F1 hybrid plants but segregates
 CC apart in subsequent generations, improving environmental safety.
 SQ Sequence 5534 BP: 1400 A: 1355 C: 1387 G: 1363 T:
 Query Match 41.1%, Score 2202, DP 25, Length 5534:
 Best Local Similarity 100.0%, Pred No 0 00e+00:
 Matches 2202; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

Db 3329 ttggacatttttgggggaattgtcggggaacccctattgtttttcttaataacatt 3388
 Cp 5360 GTGGCACTTTTGGGGGAATGTGGCGGAACCCCTATTGTGTTTATTTTCTTAATACATT 5301
 Db 5363 caaatatgatctccctcatdgaacaaataacccctgataaaagtctcaataattatgaata 3448

Cp 5300 CAATATATGATCCGCTCATGAGCAATAACCCCTGATAAATGCTTCAATATATGAAAAA 5241
 Db 3449 ggaagatgatagattcaacatttccgttgcgcttgcgcttattcccttttttgcgcaattt 3508
 Cp 5240 GGAAGAGATGAGTATCAACATTCGCTGTCGCCCTTATTCCTTTTGGGGCAATTT 5181
 Db 3509 gcttctcgttttttgcaccccaaaacgctggtgaaagttaaagatgctgaagatcaat 3568
 Cp 5180 GCTTCTCTCTTTTTCCTCACCAGAAACGCTGGTCAAAAGTAAAGATGCTGAAGATCAGT 5121
 Db 3569 tgggtcacagatgggtttacatcgaactggatctcacaacgagtaagatccttgaagatt 3628
 Cp 5120 TGGGTCCACGATGGGTTCATATCGAACTGGATCTCAACAGTGGTAAAGATCCTTGAGATT 5061
 Db 3629 ttgcgccccaaaacgcttttccaatgatgaacatttttaaagttctctctatgtggcgagg 3688
 Cp 5060 TTCGCCCCCAAGAACGCTTTTCCCAATGATGAGCACATTTTAAAGTTTCTGCIATGTGGCGCGG 5001
 Db 3689 tattatccgctattgacgccccggaagagcaactcggctgcgcgcatcacactattctcaga 3748
 Cp 5000 TATATCCCGTATTGACGCGCGGCAAGCAACTCGGTGCGCGCATACACTAATTCAGAA 4941
 Db 3749 atgacttggttgactactcaccagtcacaaacaaacatcttaccgatatgcagatcaataa 3808
 Cp 4940 ATGACTTGGTTGATCTACTCACCAGTCACAGAAAGCAATCTTACGGATGCGATGACAGTAA 4881
 Db 3809 gagaattatcagtgctgcataacccatgaatgaataacactgcgagcccaacttactctga 3868
 Cp 4880 GAGAAATTATGCACTGCTGCTGCCATAACCATGATGATGAATAACACTGCGGCCCAACTTACTCTGCA 4821
 Db 3869 caacgctcggaggaccgaaggaactaacgcctttttgcacaaacatggggaatcatgaa 3928
 Cp 4820 CAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTGTGCACAACTATGGGGATCATGTA 4761
 Db 3929 ctgccttgatcgttggaacccgagctgaatgaacccatcccaacgagcgctgaca 3988
 Cp 4760 CTCGCCCTTGATCTGTTGGGAAACCGGAGCTGAAAGGCACTACCAACGACGAGCGTGACA 4701
 Db 3989 ccacgctcgtctagcaatggcaacacgctgcgcaaaactattaaactgcgaactactta 4048
 Cp 4700 CCACGATGCTGTAGCAATGGCAACACGTTGGCGCAACTATTAACTGGCGGAACACTTAA 4641
 Db 4049 ctctagcttcgagcaacaattaatagacaggatggagcgagatgaatggagggctcagc 4109
 Cp 4640 CTCATGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAAGGACAC 4581
 Db 4109 ttctgcgctcggccttccggtgagc-ggtttatgtgtataaattcgaacccgagcagc 4168
 Cp 4580 TTCTGGCTCGGCCCTTCGGCTGCTGTTTATTGTGTATAAAATCTGGAGGCGCGTGAGC 4521
 Db 4169 gtgggtctcagctatcattgcagcactggggcgagatggtaagccctcccgctatcag 4228
 Cp 4520 GTGGGCTCTCGGGTATCATTTGAGTACTTGGGCGCAGATGGTAAAGCGCTCGGTATCGTAG 4461
 Db 4229 ttatctacacgagggagtcaggccaactatggaatgaagaaatagacaatcagcaaga 4288
 Cp 4460 TTATCTACACGACGGGAGTCAAGCAACTATGATGATGAACAGAAATAGACAGATCGCTGAGA 4401
 Db 4289 taggtcctcactgatgaagcatttgtaactcagacccaagtttactcatatacttt 4348
 Cp 4400 TAGGTGCTCTCCTTGATTAAGCATTTGTAATCTGACAGCAAGTTTACTCATATATACTTT 4341
 Db 4349 agattgatttaaaactcatttttaattaaagatatacgaagatccttttttga 4408
 Cp 4340 AGATTGATTAAACATTCATTTTAAATTTAAAGGATCTAGCTGAAGATCTTTTGTATA 4291
 Db 4409 atctcagacaaaatcccttaacgtagtttctgctcactgagcagccagccagtag 4468
 Cp 4280 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGGCGTCAGACCGCTGAG 4221
 Db 4469 aaaaatcaaaagatctcttggagatccttttttctgcgcaatactctcctctcagaaa 4528
 Cp 4220 AAAAATTAAGGATCTCTTGTAGATCTCTTTTGTGGGCTAAATCTGCTCTCTTGCATAA 4161

Wed Mar 17 09:43:34 1999

```

QY 5299 CCCCCCAATTTTGAAGCATTATCAGGGTTATTTGCTCATGAGCGGATACATATT 5298
Db 4978 tgaatgtattacaaaaataaataaataaaggttcgcgcacatttcccgaaaaagtgcc 5037
QY 5299 TGAATGTATTATGAAAAATAAACAATAGTGGTTCGCGGACATTTCCCGAAAAAGTGCC 5358
Db 5038 ac 5039
QY 5359 AC 5360

RESULT 14
ID Q40261 standard. DNA; 6926 bp.
AC Q40261;
DI C2-AUG-1993 (first entry)
DE Sequence of clone pP2-gp160MN.
KW Plasmid; cloning; restriction site; HIV-1; MN-isolate; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..3529
FT *tag= a
FT /label= pS2gpt-P2 sequences
FT 2395..2851
FT *tag= b
FT /label= E. coli 4pt gene
FT complement (3081..3423)
FT promoter
FT *tag= c
FT /label= vaccinia P7.5 promoter
FT 2395..2851
FT *tag= d
FT /label= P2 according to EP application Avipox
FT "intergenic region"
FT /note= "Oligo F artP(9)"
FT 6173..6926
FT *tag= e
FT /label= pN2gpt-S4
FT 3534..5001
FT *tag= f
FT /label= HIV-1 strain MN gp160 sequence (EMBL to
REHIVMNC)
A02211242-A
EN 04-MAR-1993
PF 25-AUG-1993; 021269
PR 26-AUG-1991; 05-750090.
PR 20-JUL-1992; 05-414736.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
PI Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PI direct molecular cloning of modified DNA molecule contg.
PI cytoplasmic DNA virus genome
PS Example: Paues 177-180; 206pp; English.
PS Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gptA and pN2-gptB which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. A new SmaI site
CC was introduced into the plasmid pN2gpt-S4 resulting in the plasmid
CC pS2gpt-S4. Subsequently the S4-promoter was exchanged by the P2-
CC promoter resulting in the plasmid pS2gpt-P2. For the modification
CC of the gp160-gene, a PCR-generated proximal fragment was exchanged
CC leading to a gp160-gene cassette with a minimal 5'-UTR.
SQ Sequence 6926 bp; 2064 A; 1470 C; 1686 G; 1706 T;

Query Match 41.0%, Score 2198; DB 7; Length 6926;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2200; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 gtggcacttttggggaaatgtgcgcgaacccctatttttttttcttaataacatt 60
QY 5360 GTGGAAATTTTGGGAAATGTTGTTGTTTATTTTCTTAATACATT 5301

```

```

CP 4220 AAAAAATCAAGAGATCTTTGACATATTTTCTTCTGAGGAAATGAGTGGTTCGAA 4161
DB 1201 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
KW 4160 AAAAAAATCAAGAGATCTTTGACATATTTTCTTCTGAGGAAATGAGTGGTTCGAA 4101
KW 1261 TCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
OS Chimeric - Mus musculus.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Bos taurus.
OS Chimeric - Rhesus macaque polyoma virus.
FH Ky Location/Qualifiers
FT Promoter 1..628
FT /*tag= a
FT /note= "CMV promoter"
FT IDNA 529..810
FT /*tag= b
FT CDS 952..1673
FT /*tag= c
FT /product= p40 subunit
FT /note= "contains an intron"
FT intron 1259..1341
FT polyA_site 1797..2024
FT /*tag= c
FT /note= "bovine growth hormone polyA site"
FT promoter 2110..2737
FT /*tag= i
FT /note= "CMV promoter"
FT IDNA 2738..2919
FT /*tag= a
FT CDS 2983..3990
FT /*tag= b
FT /product= p40 subunit
FT polyA_site 4075..4306
FT /*tag= i
FT /note= "bovine growth hormone polyA site"
PN W09745263-1.
PD 11-DEC-1997.
PR 04-JUN-1997: 009591.
PR 05-JUN-1996: 05-659206.
PA (AURA:) AURAGEN INC.
PI Rakhmillevich AL, Yang N;
DR WPI: 98-041898/04.
DR P-PSDB: W44004-05.
PT Interleukin-12 gene therapy of tumours comprises delivering
PT construct containing promoter and sequences encoding interleukin 12
PT p35 and p40 subunit(s) to target cells in vivo
PT Claim 5: Page 24-30: Supp: English.
PS Plasmid pW63196 is a tandem plasmid encoding both subunits, p35
PS p35 (see W44004) and p40 (see W44005), of murine interleukin 12
PS (IL-12). Each subunit gene was cloned from a mouse spleen cDNA
PS library and has been placed under the transcriptional control of
PS a separate cytomegalovirus (CMV) promoter. An SV40 splicing
PS donor/splicing acceptor is provided between each subunit gene and
PS its CMV promoter. The backbone of the plasmid is from pUC19.
PS pW63139 induces at least twice the expression of IL-12 as the
PS bi-directional vector pW63196 (see W6264) in vivo and in vitro. A
PS novel method of treating tumours in a mammal involves delivering
PS copies of an expressible foreign genetic construct, especially
PS pW63196 or pW63139, comprising a promoter operative in the
PS mammalian epidermal cells and RNA sequences encoding p35 and p40
PS subunits of IL-12 to target cells in vivo. Delivery of the
PS construct allows IL-12 expression for treatment of solid,
PS metastatic or disseminated tumours, and regression of established
PS tumours. The treatment is effective even when the genetic
PS construct is delivered to a site distant from the tumour.
SV Sequence 7287 BP. 1598 A, 1673 C, 1875 G, 1794 T

```

Query Match 41.0%; Score 2198; DB 49; Length 7287;

Best Local Similarity 99.9%; Prod. No. 0.000000;

Matches 2200; Conservative 0; Mismatches 2; Gaps 0;

Db 5006 gggcaatttttggggaaatgtgaggggaacccctatttttttttttaataacatt 5065
 Cp 5360 GTGGAGCTTTTGGGGGAAATGTGGCGGACACCTATTCTTTATTTTCTAAATACATT 5301
 Db 5086 caaatatataacagctatagacaataaacctcctataaatgttcaataaattgaaaaa 5125
 Cp 5300 AAAATATTATCGGCTCATGAGACAAATAACCCGATGAATGCTTCAATAATATTCAAAA 5241
 Db 5126 gaaagatgatgatattcaacatttttgggtgtggtccctttattcccttttttttgggcatatt 5185
 Cp 5240 GGAAGAGTATGAGTATTCAACATTTGGGTGTGGCTTTATTCCTTTTGGGGCATTTT 5181
 Db 5186 gcttccgttttttctcaccacgaagagctgtgaaagtataagatgctgaagatcagt 5245
 Cp 5180 GCTTTCTGCTTTTTCCTCACCACCAACAGCTGTGGAAGTAAAGATGCTGAAGATCAGT 5121
 Db 5246 tgggtgcagagtggtgtatcatcgaactggtatcacaacagcgtatgaatcttggagtt 5305
 Cp 5120 TGGTGGCAATGAGCTTACATCGAACTGGATCTCAACAGCGGTAAAGTCTTGAGATT 5061
 Db 5306 ttgcaccccaaaagcttttccaatgatgagcactttttaaagttctgctatgtggcgcg 5365
 Cp 5060 TTGGGCTGGGAAAGAGCTTTTCCAAATGATGAGCACTTTTAAAGTTCGTATGTGGGCGG 5001
 Db 5366 tatatcccgatttgaacgcgggcaagagcaactcgttcgcgcgatacacattctcaga 5425
 Cp 5000 TATTATCCGATATGAGCGGGGCAAGAGCAACTCGGTTCGGCGCATACACTATTCTCAGA 4941
 Db 5426 ataatgtgtgaagtactcaacagcaagaaagcactttacggatggcatcacagtaa 5485
 Cp 4940 ATGACTTGGTGTGASTACTCACCAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAA 4881
 Db 5486 quaatattacagttcgtcgaataaacctdggtgatcaactggtggcgcgaantttacttga 5545
 Cp 4880 GAGAAATTATCAGTGTGCTGCATTAACCATGAGTGAACACTGCGGCCAACTTACTTCTGA 4821
 Db 5546 caacgatcgaagagagagagagctaacgctttttgacacaacatggggatcatataa 5605
 Cp 4820 CAACGATCGGAGGACCGAAGAGCTTAACCGCTTTTTCGACACATGGGGATCATGTAA 4761
 Db 5606 ctgagcttgaatcgtttggaacccggaactgaaatgaagcctatacaaacagagagctga 5665
 Cp 4760 CTGCGCTTGAATGCTTGGAGCGGAGCTGAATGAAGCACTAACCAACGACGAGGTGACA 4701
 Db 5666 caagcagatcgtatgaatgaagagagagagctaacgctttttgacacaacatggggatcatataa 5725
 Cp 4700 CCAATGATGCTGTATGCAATGGCAACAGCTTGGCGAAACTTATTAAGTGGCGNACTACTTA 4641
 Db 5726 ctctagcttcggggaacaaatgaatgaactggatggagcgaataaaadtgcaggaccac 5785
 Cp 4640 CTCTAGCTTCTGGGCAACATTAATAGACTGGATGGAGCGGATAAAGTTGCGAGACAC 4581
 Db 5786 ttctgctgttggccttccggctgagctggttttattctgataaatacgaagcggatgagc 5845
 Cp 4580 TTCTGCTGTGGGCTTCCGGCTGGCTGGTTAATGATGATGAGGCGGTGAGC 4521
 Db 5846 atagatcgtgggtatgattgagcagctggggcgaatgataagcctccctcctatcgtag 5905
 Cp 4520 GTGGGCTGTGCTGATCATGATGACGCTGGGGCAGATGCTAAGGCTTCCGCTATCTAG 4461
 Db 5906 ttatctacacgagggggggagtcagggaactatgagagacgaataacagatcgtgaga 5965
 Cp 4460 TTAATGATGAGGAGGAGTCTAGGCACTATGGATGAGCAATAGACAGATGCTGTAGA 4401
 Db 5966 tangagctcagtgattgaagattgaactgtcagaccgaagtttactcatatatacttt 6025
 Cp 4400 TAGTGCTCTACTGATTAAAGCATTTGGTAATGCTCAGACCAAGTTTACTCATATATACTTT 4341
 Db 6026 aqatttgatttaaaccttcttttaatttaaaagagatgaagagatcctttttcata 6085
 Cp 4340 AGAATGATTAAACTTCAATTTTAATTTAAAGGATCTAGGTGAAGATCTTTTGTGATA 4281
 Db 6086 atctcatgaccacaaatcccttaacgtgagttttcttccactgagctcagaccgctag 6145

Cp 4280 ATCTCATGACCAAAATCCCTTAAGCTGAGTTTCSTTCCACTGAGTGTGAGAGCTGTGAG 4221
 Db 6146 aaaaatcaaaagatcctcttggatcctcttttttctggcgtaactgtgctgctggcaaa 6205
 Cp 4220 AAAAGATCAAAAGATCTCTTGAGATCCTTTTCTGGCGGTAACTGCTGCTTGCATA 4161
 Db 6206 caaaaaaacccaccctaccagcgtgtgtgtgtgtgcccggatcaagatcaacatcttt 6265
 Cp 4160 CAAAAAACCCCGCTACGCGCTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTT 4101
 Db 6266 ttccgaagtaactgcttcagcagcgcagatcaccaatctatgctccttctagttagc 6325
 Cp 4100 TTCCGAAGTAACTGGCTTCAGCAGAGCGGAGATACCAATACTGTCTTCTAGTGTAGC 4041
 Db 6326 cgtagttaggcacacacttcaagadctctgttagcaccgcctacatcctcgtctgctaa 6385
 Cp 4040 GGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACTGTAGCACTGCTGCTGTAA 3981
 Db 6386 tectgttaccagtgcctcctccagtgcagataagtcgtgtctttaccgggttggactcaa 6445
 Cp 3980 TCTGTTCACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTGTCTTACGCGTCA 3921
 Db 6446 gacgatgttaccggataaaggcgcagcgtgggctcaacgggqddtctgtacacacac 6505
 Cp 3920 GAGCATATACGGATAAGCGCAGCGTGGGCTGAACGGGGGTTCGTGCACACAGC 3861
 Db 6506 coagcttgagcgaacacactacacacgaactgaatacctacagcgtgagctatgaaaaa 6565
 Cp 3860 CCAGCTTGGAGCAAGACCTTACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA 3801
 Db 6566 gcgcacgcttcccgagggagaaaggcggcagagataccgystaaggagagagctcgaag 6625
 Cp 3800 GCGCACGCTTCCGGAAGGAGAAAGCGCACAGATATCCGCTAAGCGGAGGCTCGAAA 3741
 Db 6626 cagagagcagcagcagggagcttccaaaggaaacgcctgctattcttatagctgtcgc 6685
 Cp 3740 CAGAGAGCGCACAGAGGACCTTCCAGGGGGAACCGCTTGTATGCTTATAGTGTG 3681
 Db 6686 ggttctgcacactcgtgacttgagctcgatctttgtgagctcgtcaggggggcgcagcc 6745
 Cp 3680 GGTITCGCCACTCTGACTTGGCTGATTTTGTGATGCTCGTCAAGGGGCTGAGGCC 3621
 Db 6746 tatggaaaaacgcagcaaacgcgccttttaccggttctcctgccccttttggcttttg 6805
 Cp 3620 TATGMAAAACCGCAGCAACCGGCTTTTACGCTTCTGTGGGCTTATGCTTGGGCTT 3561
 Db 6806 ctccatgcttcttctgctgttatccctgattctgtgataacccctattacccgctttg 6865
 Cp 3560 CTCACATGTTCTTCTCTGCTGCTATCTCCCTGATTCTGTGATAACCGTATTACGGCTT 3501
 Db 6866 agtgagctgatacgcctcgc 6925
 Cp 3500 AGTGTGCTGATACCTGCTCGCGCAGCGCAAGCGACGAGCGCAGCGAGTCACTGAGT 3441
 Db 6926 aagcgaagagc 6985
 Cp 3440 AAGCGGAAGAGCGCCCATACAGCAAAACGCTCTCTCCCGCGCTTGGCGGATTCAITAT 3381
 Db 6986 cgagctggcagcagcaggggttcccaacagcgaagcgaagcgaagcgaagcgaagcgaag 7045
 Cp 3380 GCAGCTGGCAGCAGCAAGTTTCCGACTGGAAAGCGGCGAGTGAGCGCAACATTAATG 3321
 Db 7046 tgaattagctcctcctattaggaaccccgagctttacactttatgcttccggcctgctatgt 7105
 Cp 3320 TGAGTTAGCTCAGTCACTATAGGACACCCAGCTTTTACACTTTTATGCTTCCGGCTG 3261
 Db 7106 tgtgtggaattgtgagcggatcaacatttccacacaggaacacagctatgacatgcttacg 7165
 Cp 3260 TGTGTGGAATTGTGAGCGGATAACCAATTCACACAGGAAACAGCTATGAGCAACCAAT 3201
 Db 7166 ccaagctcgaataaaccctcactcaaaagggaacaaacacacacacacacacacacac 7207

CP 4200 00AAGGCGCAATTACGCTTACTAAAGCGGACAAAGCTGG 4159

Search completed: Sun Mar 14 23:33:08 1999
Job time : 1057 secs.

MPSEARCH

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Sun Mar 14 20:54:50 1999; MasPar time 6810.13 Seconds
1409.454 Million cell updates/sec
Tabular output not generated.

Title: >US-09-020-716-3
Description: (1-5360) from US09-020-716 seq
Perfect Score: 5360
N.A. Sequence: 1 CTAAATGTAAGCTGTTAATAATTTCGCCGAAAGTGCAC 5360
Comp: JATTTAACATTCGCATATTAAAGGGGCTTTTCACGGTG

Scoring table:
Gap 6
TABLE default

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est56
1:em_est1 2:em_gss1 3:em_gss2 4:em_gss3
Database: genbank-est109
5:gb_est1 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13
10:gb_est14 11:gb_est15 12:gb_est16 13:gb_est17
14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21
18:gb_est22 19:gb_est23 20:gb_est24 21:gb_est25 22:gb_est26
23:gb_est27 24:gb_est28 25:gb_est29 26:gb_gss1 27:gb_gss2
28:gb_gss3 29:gb_gss4

Statistics: Mean 12.787; Variance 2.653; scale 4.820

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	654	12.4	711	26	AG004608 Homo sapiens genomic D	0.00e+00
2	654	12.4	711	27	AG010947 Homo sapiens genomic D	0.00e+00
3	657	12.3	721	27	AG011001 Homo sapiens genomic D	0.00e+00
4	657	12.3	721	26	AG004662 Homo sapiens genomic D	0.00e+00
5	630	11.8	698	26	AG003787 Homo sapiens genomic D	0.00e+00
6	630	11.8	698	27	AG009976 Homo sapiens genomic D	0.00e+00
7	635	11.8	718	26	AG004363 Homo sapiens genomic D	0.00e+00
8	635	11.8	718	27	AG010489 Homo sapiens genomic D	0.00e+00
9	624	11.5	696	26	AG009755 Homo sapiens genomic D	0.00e+00
10	624	11.5	720	26	AG013858 Homo sapiens genomic D	0.00e+00
11	624	11.5	720	26	AG005062 Homo sapiens genomic D	0.00e+00
12	624	11.5	720	26	AG005062 Homo sapiens genomic D	0.00e+00
13	619	11.5	747	26	AG007052 Homo sapiens genomic D	0.00e+00

C	14	609	11.4	644	28	AQ115013	CIT-HSP-2324C24.TF.CIT	0.00e+00
C	15	603	11.3	692	27	AQ040352	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	16	603	11.3	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	598	11.2	748	17	A1124281	1-53 Drosophila 8-12.H	0.00e+00
C	18	593	11.1	640	27	AQ074298	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	19	585	11.1	654	26	B78704	CIT978SK-A-880F6.TPC.C	0.00e+00
C	20	594	11.1	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	21	597	11.1	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	22	588	11.0	652	28	A1108633	CIT-HSP-2327M3.TF.CIT	0.00e+00
C	23	583	10.9	594	26	C17B6	C10a intestinalis gen	0.00e+00
C	24	583	10.9	666	27	AQ079096	CIT-HSP-2355E24.TF.CIT	0.00e+00
C	25	577	10.8	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.6	642	28	AQ113213	CIT-HSP-2372L22.TF.CIT	0.00e+00
C	27	566	10.6	682	27	AQ074693	CIT-HSP-2376F2.TF.CIT	0.00e+00
C	28	557	10.4	615	28	AQ12424	CIT-HSP-2383F2.TF.CIT	0.00e+00
C	29	559	10.4	742	28	AQ108392	CIT-HSP-2383F2.TF.CIT	0.00e+00
C	30	551	10.3	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.3	634	27	AQ038010	CIT-HSP-2385K21.TF.CIT	0.00e+00
C	32	549	10.2	653	27	AQ077504	CIT-HSP-2345M2.TF.CIT	0.00e+00
C	33	545	10.2	661	27	AQ057603	PPC111-19F1.TP.RPC11	0.00e+00
C	34	545	10.2	690	27	AQ029299	CIT-HSP-2336K22.TF.CIT	0.00e+00
C	35	544	10.1	633	27	AQ040787	CIT-HSP-2365E1.TF.CIT	0.00e+00
C	36	540	10.1	637	27	AQ077360	RPC111-9N1.TP.RPC11.H	0.00e+00
C	37	534	10.0	594	26	B71605	CIT-HSP-2371M2.TF.CIT	0.00e+00
C	38	536	10.0	626	28	AQ111342	PPC111-19F5.TPB.RPC11	0.00e+00
C	39	531	9.9	549	27	AQ003191	PPC111-19113.TPB.RPC11	0.00e+00
C	40	530	9.9	630	27	AQ003238	Homo sapiens genomic D	0.00e+00
C	41	526	9.8	743	26	AG002826	CIT-HSP-2382N1.TF.CIT	0.00e+00
C	42	526	9.8	540	27	AQ074307	Homo sapiens genomic D	0.00e+00
C	43	526	9.8	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C	44	525	9.8	600	28	AQ111411	CIT-HSP-2373F2.TF.CIT	0.00e+00
C	45	527	9.8	629	27	AG011357	Homo sapiens genomic D	0.00e+00

ALIGNMENTS

RESULT 1 AG004608 711 bp DNA GSS 29-JAN-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone: S594B638 genomic
DEFINITION survey sequence.
ACCESSION AG004608
NID: q2822058
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: S594B638.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates, Catarrhini, Hominoidea, Homo
REFERENCE 1 (bases 1 to 711)
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
JOURNAL Published Only in Database (1998) In press
TITLE Homo sapiens genomic DNA, chromosome 21q
REFERENCE 2 (bases 1 to 711)
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
JOURNAL Direct Submission
TITLE Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.
JOURNAL Masahira Hattori, Kitasato University, Department of Science, JSI
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@jsgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)

FEATURES
source
Location/Qualifiers
1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="S594B638"
/map="21q"
BASE COUNT 169 a 188 c 172 g 182 t

ORIGIN
Query Match 12.4%; Score 664; DB 26; Length 711;
Best Local Similarity 99.6%; Prog No 0; Gaps 3;
Matches 682; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

```

DB 28 TATATAATACGATCACTGAGAGAGTTTATCACTGGGCTGCAAGTGGTAAATACATACGCG 87
QY 4454 TATATAATACGATCACTGAGAGAGTTTATCACTGGGCTGCAAGTGGTAAATACATACGCG 4512
DB 88 AGAGTACAGTCTACGCGCTGACAGATTATCAGCAATAAAGAGAGAGAGAGAGAGAGAG 147
QY 4513 AGAGTACAGTCTACGCGCTGACAGATTATCAGCAATAAAGAGAGAGAGAGAGAGAGAG 4572
DB 148 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
QY 4573 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4632
DB 208 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 267
QY 4633 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4692
DB 268 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
QY 4693 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4752
DB 328 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY 4753 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4812
DB 388 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
QY 4813 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4872
DB 448 TAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 507
QY 4873 TAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4932
DB 508 CAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 567
QY 4933 CAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4992
DB 568 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
QY 4943 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5052
DB 628 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686
QY 5053 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5112
DB 687 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
QY 5112 GTGACCCCAACTGATCTTCAGATC 5136

PEST 2 AG010947 711 bp DNA GSS 28-JUL-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 8594R38, genome survey sequence.
ACCESSION AG010947
KEYWORDS g334775
SOURCE GSS.
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 711)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Homo sapiens genomic DNA, chromosome 21q
Published only in Database (1998) In press
REFERENCE 2 (bases 1 to 711)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Direct Submission
Submitted (28-JUL-1998) to the DDBJ/EMBL/GenBank databases
Masahiro Hattori, Kitasato University, Department of Science, 1ST
Symposium Laboratory, Kitasato University, Sagamihara 228, Japan
(E-mail:hattori@life.kitakyu.ac.jp Tel: 0427-78-9932
Fax: 0427-78-9561)
Location/Qualifiers
FEATURES

```

```

source 1..711
/organism="Homo sapiens"
/locus="xref=*,axon:9604"
/chromosome="21"
/clone="S594R38"
/map="21q"
BASE COUNT 159 a 188 c 172 g 192 t
ORIGIN
Query Match 12.4% Score 664; DB 27; Length 711;
Best local similarity 99.68; Pred. No. 0.00e+00;
Matches 682; Conservative 0; Mismatches 6; Indels 0; Gaps 0
DB 28 TATATAATACGATCACTGAGAGAGTTTATCACTGGGCTGCAAGTGGTAAATACATACGCG 87
QY 4454 TATATAATACGATCACTGAGAGAGTTTATCACTGGGCTGCAAGTGGTAAATACATACGCG 4512
DB 88 AGAGTACAGTCTACGCGCTGACAGATTATCAGCAATAAAGAGAGAGAGAGAGAGAGAGAG 147
QY 4513 AGAGTACAGTCTACGCGCTGACAGATTATCAGCAATAAAGAGAGAGAGAGAGAGAGAGAG 4572
DB 148 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
QY 4573 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4632
DB 208 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 267
QY 4633 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4692
DB 268 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327
QY 4693 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4752
DB 328 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY 4753 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4812
DB 388 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 447
QY 4813 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4872
DB 448 TAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 507
QY 4873 TAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4932
DB 508 CAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 567
QY 4933 CAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4992
DB 568 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
QY 4943 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5052
DB 628 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686
QY 5053 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5112
DB 687 GAGTAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
QY 5112 GTGACCCCAACTGATCTTCAGATC 5136

RESULT 3 AG011001
LOCUS Homo sapiens genomic DNA, 21q region, clone: 879612X91, genome survey sequence.
ACCESSION AG011001
KEYWORDS g3347935
SOURCE GSS.
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) in press
REFERENCE 2 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
Location/Qualifiers
1..721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="879G12X91"
/map="21q"
BASE COUNT 195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 12.3%, Score 657, DB 27, Length 721;
Best Local Similarity 98.7%; Pred.No.0.00e-00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATATGCGAGTGGTGGCATACCATGATGAT 91
Cp 4905 CATCTTACGGATGGCATGACAGTAAGAGAAATATGCGAGTGGTGGCATACCATGATGAT 4846
Db 92 AACACTGGCGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 151
Cp 4945 AACACTGGCGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 4786
Db 152 TTGCACAACTGGCGGATCATATGTAACCTGCGCTTGATCGTGGTGGTGGTGGTGGT 211
Cp 4785 TTGCACAACTGGCGGATCATATGTAACCTGCGCTTGATCGTGGTGGTGGTGGTGGT 4726
Db 212 GGCATACCAAAAG 271
Cp 4725 GGCATACCAAAAG 4666
Db 272 AACACTTAATTCGGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 331
Cp 4845 AACACTTAATTCGGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 4796
Db 332 GAGGCGGATTAAGATTCGAGGACCACTTCTGCGGCTGGGCTGGGCTGGGCTGGGCT 391
Cp 4605 GAGGCGGATTAAGATTCGAGGACCACTTCTGCGGCTGGGCTGGGCTGGGCTGGGCT 4546
Db 442 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 511
Cp 4425 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4486
Db 452 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 511
Cp 4485 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4426
Db 512 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 570
Cp 4425 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4366
Db 571 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 629
Cp 4365 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4306
Db 629 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 687
Cp 4305 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4246
Db 687 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 719
Cp 4245 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4214

RESULT 4
LOCUS AG004662 721 bp DNA GSS 20-JAN-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic survey sequence.
ACCESSION AG004662
NID 52826191
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 879G12X91.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 721)
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) in press
REFERENCE 2 (bases 1 to 721)
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
Location/Qualifiers
1..721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="879G12X91"
/map="21q"
BASE COUNT 195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 12.3%, Score 657, DB 26, Length 721;
Best Local Similarity 98.7%; Pred.No.0.00e-00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATATGCGAGTGGTGGCATACCATGATGAT 91
Cp 4905 CATCTTACGGATGGCATGACAGTAAGAGAAATATGCGAGTGGTGGCATACCATGATGAT 4846
Db 92 AACACTGGCGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 151
Cp 4845 AACACTGGCGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 4796
Db 152 TTGCACAACTGGCGGATCATATGTAACCTGCGCTTGATCGTGGTGGTGGTGGTGGT 211
Cp 4785 TTGCACAACTGGCGGATCATATGTAACCTGCGCTTGATCGTGGTGGTGGTGGTGGT 4726
Db 212 GGCATACCAAAAG 271
Cp 4725 GGCATACCAAAAG 4666
Db 272 AACACTTAATTCGGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 331
Cp 4845 AACACTTAATTCGGGCAACTTACTTCTGACACAGTCCGAGGACCGAAGGAGCTAACCGCTTTT 4796
Db 332 GAGGCGGATTAAGATTCGAGGACCACTTCTGCGGCTGGGCTGGGCTGGGCTGGGCT 391
Cp 4605 GAGGCGGATTAAGATTCGAGGACCACTTCTGCGGCTGGGCTGGGCTGGGCTGGGCT 4546
Db 442 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 511
Cp 4425 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4486
Db 452 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 511
Cp 4485 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4426
Db 512 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 570
Cp 4425 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4366
Db 571 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 629
Cp 4365 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4306
Db 629 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 687
Cp 4305 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4246
Db 687 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 719
Cp 4245 GATGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4214

Wed Mar 17 09:43:39 1999

Db 151 CACAAATGGGGGATCATCTAACTGCGCTTCATGCTTGGGAACCGGAGCTGAATGAAGCC 210
 Cp 4782 CACAAATGGGGGATCATCTAACTGCGCTTCATGCTTGGGAACCGGAGCTGAATGAAGCC 4723
 Db 211 ATACZAAATGAGAGGCTGACACTACGATGCTGCTATCAATAGGAAATGAGCTTGGCAAA 270
 Cp 4722 ATACZAAATGAGAGGCTGACACTACGATGCTGCTATCAATAGGAAATGAGCTTGGCAAA 4663
 Db 271 CTATTAAGTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 330
 Cp 4662 CTATTAAGTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 4603
 Db 331 GCGAATAAGCTTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 380
 Cp 4602 GCGAATAAGCTTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 4543
 Db 381 GATAAATCTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 450
 Cp 4542 GATAAATCTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 4483
 Db 451 GATAAATCTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 509
 Cp 4482 GATAAATCTGGGCAATCTTACTCTAGCTTCCCGGGAACCAATTAATAGACTGGATGGAG 4423
 Db 510 GAAATAGACAGATGCTGAGATAGTGGCTCACTGATTAGCATTTGGTAACTGTCAGA 569
 Cp 4422 GAAATAGACAGATGCTGAGATAGTGGCTCACTGATTAGCATTTGGTAACTGTCAGA 4364
 Db 570 CCAATTTACTGCTATATATTTAGATTTGATTTAAAGCTTCAATTTTAAATTTAAAGGAT 628
 Cp 4363 CCAATTTACTGCTATATATTTAGATTTGATTTAAAGCTTCAATTTTAAATTTAAAGGAT 4304
 Db 629 CTAGGCAAGATTTTATGATATCTATGACCAATTTCCCTTAACGTTGATTTTGGTT 687
 Cp 4303 CTAGGCAAGATTTTATGATATCTATGACCAATTTCCCTTAACGTTGATTTTGGTT 4244
 Db 688 TCACATGA 695
 Cp 4243 TCACATGA 4237
 RESULT 7 AG004363 718 bp DNA GSS 17-JAN-1998
 LOCUS Homo sapiens genomic DNA, 21q region, clone: P7H8X25, genomic
 DEFINITION survey sequence
 ACCESSION AG004363
 NID g2789511
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: P7H8X25.
 ORGANISM
 Eukaryotes: Chordata; Vertebrata; Mammalia; Eutheria;
 Primates: Catarrhini; Hominidae; Homo.
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 718)
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (17-JAN-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
 (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, tel: 0427-78-9732,
 Fax: 0427-78-9561)
 Location/Qualifiers
 1..718
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21q"
 /clone="P7H8X25"
 /map="21q"
 FEATURES
 source
 BASE COUNT 176 a 191 c 187 g 158 t 6 others
 ORIGIN

Query Match 11.8%; Score 635; DR 26; Length 718;
 Best Local Similarity 97.3%; Pred. No. 0.00e+00;
 Matches 682; Conservative 0; Mismatches 13; Indels 6; Gaps 5;
 Db 21 TCACAGAAATCAAGGATACGAGCAAGAACATGTGAGCAAAAGGCTGACAAAGGCT 80
 Cp 3521 TCACAGAAATCAAGGATACGAGCAAGAACATGTGAGCAAAAGGCTGACAAAGGCT 3580
 Db 81 AGAAGGCTTAAAGGCTGAGGCTGAGGCTTTTCCATAGGCTCCGCTCCCTGAGCAG 140
 Cp 3581 AGAAGGCTTAAAGGCTGAGGCTGAGGCTTTTCCATAGGCTCCGCTCCCTGAGCAG 3640
 Db 141 CATACAAAAATGAGGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCA 200
 Cp 3641 CATACAAAAATGAGGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCA 3700
 Db 201 CAGGCTTTCCGCTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 260
 Cp 3701 CAGGCTTTCCGCTTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 3760
 Db 261 GATACCTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTT 320
 Cp 3761 GATACCTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTT 3820
 Db 321 AGGTATCTCAGTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 380
 Cp 3821 AGGTATCTCAGTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 3880
 Db 381 GTTACGCTGAGGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 440
 Cp 3881 GTTACGCTGAGGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 3940
 Db 441 CAGCATTATCCGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 500
 Cp 3941 CAGCATTATCCGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 4000
 Db 501 GGGGTGCTTAAAGGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 560
 Cp 4001 GGGGTGCTTAAAGGCTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 4058
 Db 561 TATTGGTATCTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 620
 Cp 4059 TATTGGTATCTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCTTTTCCGCT 4117
 Db 621 TGATCGGGCAAAACCAAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 679
 Cp 4118 TGATCGGGCAAAACCAAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 4177
 Db 680 ACGCGCAGAAAAAAGGATCTCAAGA-GATC-TTTCATCTT 718
 Cp 4178 ACGCGCAGAAAAAAGGATCTCAAGA-GATC-TTTCATCTT 4218
 RESULT 8 AG010489 718 bp DNA GSS 16-JUL-1998
 LOCUS Homo sapiens genomic DNA, 21q region, clone: P7H8X25, genomic
 DEFINITION survey sequence.
 ACCESSION AG010489
 NID g3327308
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: P7H8X25.
 ORGANISM
 Eukaryotes: Chordata; Vertebrata; Mammalia; Eutheria;
 Primates: Catarrhini; Hominidae; Homo.
 Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published Only in Database (1998) In press
 REFERENCE 2 (bases 1 to 718)
 Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-1998) to the DDBJ/EMBL/GenBank databases.


```

Cp 4575 GATCGTTGGAAACGGAGCTGAATGAAGCCATACCAACACAGAGCGTGA-CACCAAGA: 4594
Db 640 GATCTC 645
Cp 4517 GATCTC 4512

RESULT 10
LOCUS AG003576 696 bp DNA GSS 19-DEC-1997
DEFINITION Homo sapiens genomic DNA, 21q region, clone: P8G4SpN15, genomic
survey sequence.
ACCESSION AG003576
NID 92705702
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P8G4SpN15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in DataBase (1997) In press
REFERENCE 2 (bases 1 to 696)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagami-hara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
source Location/Qualifiers
1..696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P8G4SpN15"
/map="21q"
BASE COUNT 183 a 165 c 185 g 162 t 1 others
ORIGIN
Query Match 11.6%; Score 624; DB 26; Length 696;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 650; Conservative 0; Mismatches 0; Indels 6; Gaps 6;

Db 31 GTTTTGTCTACCCAGAACGCTGGTGAAGCTAAAGATGCTGAAGATCAGTTGGGTGCA 90
|||||
Cp 5172 GTTTTGTCTACCCAGAACGCTGGTGAAGCTAAAGATGCTGAAGATCAGTTGGGTGCA 5113
|||||
Db 91 CGAGTGGGTTAGATCGAACTGGATCTCAACAGCGGTAAGATCCCTTGAGAGTTTTCGCCCC 150
|||||
Cp 5112 CGAGTGGGTTAGATCGAACTGGATCTCAACAGCGGTAAGATCCCTTGAGAGTTTTCGCCCC 5053
|||||
Db 151 GAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTCTGCTATGTCGGCGGGTATTATCC 210
|||||
Cp 5052 GAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTCTGCTATGTCGGCGGGTATTATCC 4993
|||||
Db 211 GATTAATGAGGAGGAGAGAGAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 270
|||||
Cp 4992 GATTAATGAGGAGGAGAGAGAGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4933
|||||
Db 271 GTTGATATCTACAGCTGACAGAAAGATCTTTAGGAGATGCATGACAGTAGAAGATTATTA 330
|||||
Cp 4932 GTTGATATCTACAGCTGACAGAAAGATCTTTAGGAGATGCATGACAGTAGAAGATTATTA 4873
|||||
Db 331 TCGAGTGGTGGCATACCACTGATGATACACTGCGGCGCAACTTCTCTGACACAGCATC 390
|||||
Cp 4872 TCGAGTGGTGGCATACCACTGATGATACACTGCGGCGCAACTTCTCTGACACAGCATC 4813
|||||
Db 391 GGAAGAGCGAAAGAGTAACTGGCTTTTGGCACAACATGCGGGGATCATGTAACTGGCCTT 450
|||||
Cp 4812 GGAAGAGCGAAAGAGTAACTGGCTTTTGGCACAACATGCGGGGATCATGTAACTGGCCTT 4753
|||||
Db 451 GATCGTTGGAAACGGAGCTGAATGAAGCCATACCAACACAGAGCGTGA-CACCAAGA: 4594

```

```

Cp 4752 GATCGTTGGAAACGGAGCTGAATGAAGCCATACCAACACAGAGCGTGA-CACCAAGA: 4594
Db 511 GGCCTGTAGCAAGGCAACAGCTTGGCGAACTATTAACTGGCGAACTACTACTCTAG 570
|||||
Cp 4693 G-CCTGTAGCAATGCAACAGCTTGGCGAACTATTAACTGGCGAACTACTACTCTAG 4635
|||||
Db 571 CTTCCCGGCAACAATTAATAGATGGATGAGGCGGATAAAGTTGCGAGGACCACTTCTG 630
|||||
Cp 4634 CTTCCCGG-CAACAATTAATAGATGGATGAGGCGGATAAAGTTGCGAGGACCACTTCTG 4576
|||||
Db 631 CGCTGGCCCTTTCGGCTGGCTGCTTTATTGCTGATAAATCTGAGCGCGGTGAGC-TG 689
|||||
Cp 4575 CGCTGGCCCTT-CCGGCTGG-CTGCTTTATTGCTGATAAATCTGAGCGCGGTGAGC-TG 4518
|||||
Db 690 GGTCTC 695
Cp 4517 GGTCTC 4512

RESULT 11
LOCUS AG013858 720 bp DNA GSS 10-SEP-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic
survey sequence.
ACCESSION AG013858
NID 93560348
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 762015SpN14.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in DataBase (1998) In press
REFERENCE 2 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagami-hara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
source Location/Qualifiers
1..720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="762015SpN14"
/map="21q"
BASE COUNT 172 a 178 c 170 g 186 t 14 others
ORIGIN
Query Match 11.6%; Score 624; DB 28; Length 720;
Best Local Similarity 97.2%; Pred. No. 0.00e+00;
Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGCTCCCTCGTGGCGCTCTCTCTGTTCCACGCTGTCGCTTAACTGGAATCTTGGCGCTT 94
|||||
Cp 3719 AGCTCCCTCGTGGCGCTCTCTCTGTTCCACGCTGTCGCTTAACTGGAATCTTGGCGCTT 3779
|||||
Db 95 CTCCTCTCGGGAAGCGTGGCGCTTCTCATAGTCTCAAGCTGTAAGTATCTCAGTTGGTG 154
|||||
Cp 3779 CTCCTCTCGGGAAGCGTGGCGCTTCTCATAGTCTCAAGCTGTAAGTATCTCAGTTGGTG 3838
|||||
Db 155 TAGTGTGTTGGCTCCAAAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 214
|||||
Cp 3839 TAGTGTGTTGGCTCCAAAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3838
|||||
Db 215 GCTTATCCGCTGAATATCTGCTTGTAGTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 274
|||||
Cp 3899 GCTTATCCGCTGAATATCTGCTTGTAGTCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3958
|||||

```


Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
7712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@igir.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
Location/Qualifiers
1..644

/organism="Homo sapiens"
/note="vector: pBelBAC11; Site_1: HindIII; Site_2:
HindIII"
/db_xref="taxon:9606"
/clone="2374C24"
/clone_lib="Cif-Hsp"
/sex="Male"
/cell_type="Sperm"
135 a 183 c 145 g 179 t 2 others

BASE COUNT 135 a 183 c 145 g 179 t 2 others
ORIGIN

Query Match 11.4%; Score 609; DB 28; Length 644;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

5 TCACTGGCGCTGCTTTTACAGCTGCTGAGTGGGAAACDCTGGCGTTACCCAACTTAAT 64
Cp 618 TCATGGCGCTGCTTTTACAGCTGCTGAGTGGGAAACDCTGGCGTTACCCAACTTAAT 559
Db 65 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAAATAGGAGAGAGAGAGAGAG 124
Cp 558 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAAATAGGAGAGAGAGAGAGAG 499
Db 125 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAAATAGGAGAGAGAGAGAGAG 193
Cp 498 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAAATAGGAGAGAGAGAGAGAG 439
Db 184 TTAAGCGGGGGGGGGTGGTGGTTACGCCAGCTGACGCTACACCTGCGAGAGAGAGAG 243
Cp 438 TTAAGCGGGGGGGGGTGGTGGTTACGCCAGCTGACGCTACACCTGCGAGAGAGAGAG 379
Db 244 GCGCGCGCTGCTTCGCTTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 303
Cp 378 GCGCGCGCTGCTTCGCTTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 319
Db 304 CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTAGTGGTTTACGGAGAGAGAGAG 463
Cp 318 CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTAGTGGTTTACGGAGAGAGAGAG 259
Db 364 CCCAAACAACTTGATTAGGGTGGTGGTTACGGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Cp 259 CCCAAACAACTTGATTAGGGTGGTGGTTACGGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
Db 424 TTTCGCCCTTTACGCTTGGAGTCCAGCTTCCTTTAAATAGTGGAGCTCTTTTCCAACTTGA 483
Cp 198 TTTCGCCCTTTACGCTTGGAGTCCAGCTTCCTTTAAATAGTGGAGCTCTTTTCCAACTTGA 139
Db 484 ACAACACTCAACCTATCTCGGCTCTATCTTTTGGATTATTAAGGGATTTCGGCGATTTCG 542
Cp 138 ACAACACTCAACCTATCTCGGCTCTATCTTTTGGATTATTAAGGGATTTCGGCGATTTCG 79
Db 544 GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTAACGGCAATTTTACCAAAATA 603
Cp 78 GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTAACGGCAATTTTACCAAAATA 19

TITLE
JOURNAL
COMMENT

BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 11.5%; Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

2b 41 ATCTGGCTCTGCTGAGGAGGAGTACCTTGGGAAAGAGTTGGTAGCTCTTGATCCGGC 100
Qy 4057 ATCTGGCTCTGCTGAGGAGGAGTACCTTGGGAAAGAGTTGGTAGCTCTTGATCCGGC 4126
Db 101 AAACAACAACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 159
Qy 4127 AAACAACAACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4186
Db 160 AAAAAGAGATCTCAAGAGATCTCTTTGATCTTTTACGGGGTCTGACGCTCAGTGGAAAC 219
Qy 4187 AAAAAGAGATCTCAAGAGATCTCTTTGATCTTTTACGGGGTCTGACGCTCAGTGGAAAC 4246
Db 220 GAAAGACTCAGCTTAAGGAGTTTGGTCATGAGATTATCAAAAGAGTCTTACCTAGATC 279
Qy 4247 GAAAGACTCAGCTTAAGGAGTTTGGTCATGAGATTATCAAAAGAGTCTTACCTAGATC 4306
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT 339
Qy 4307 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT 4366
Db 340 GACAGTTACCAATGCTTAATCAGTGGGAGGAGTCTTACGAGGAGTCTTCTTATTCGTTCA 399
Qy 4367 GACAGTTACCAATGCTTAATCAGTGGGAGGAGTCTTACGAGGAGTCTTCTTATTCGTTCA 4426
Db 400 TCATAGTGGCTGAGTCCCGCTGCTGTAGATACTACATACATACATACATACATACATACAT 458
Qy 4427 TCATAGTGGCTGAGTCCCGCTGCTGTAGATACTACATACATACATACATACATACATACAT 4486
Db 459 GAGGCTAGTCTCAATGATACGCGGAGACCCACGCTCAGCGGCTCCAGATTATACAGA 518
Qy 4487 GAGGCTAGTCTCAATGATACGCGGAGACCCACGCTCAGCGGCTCCAGATTATACAGA 4546
Db 519 ATAAACAGGAG 577
Qy 4547 ATAAACAGGAG 4605
Db 578 CATCCAGTCTATTAATTTGTCGGGAGAGCTAGAGTAAGTAAAGTATTCGCCAGTTAATAGT 637
Qy 4606 CATCCAGTCTATTAATTTGTCGGGAGAGCTAGAGTAAGTAAAGTATTCGCCAGTTAATAGT 4663
Db 638 TTGGGAGAGCTGTTGNCATTTGTTGAGGAGATCCTGCTGAGGAGGAGGAGGAGGAGGAG 697
Qy 4664 TTGGGAGAGCTGTTGNCATTTGTTGAGGAGATCCTGCTGAGGAGGAGGAGGAGGAGGAG 4722
Db 698 GTCTTCATTCAGTTCGGGTTCCCAACGATCAAGGGGGTTACATGAT 746
Qy 4723 GTCTTCATTCAGTTCGGGTTCCCAACGATCAAGGGGGTTACATGAT 4769

RESULT 14
LOCUS A0115013 644 bp DNA GSS 28-AUG-1998
DEFINITION Cif-HSP-2374C24.TP Cif-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION A0115013
NID 33491134
KEYWORDS GSS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 644)
Adams, M.D., Founley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and

W P S R E F

(TW)

Release 3 JA John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a - n.a database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 01:29:29 1999. Maspar time 8427.16 Seconds
1545.419 Million cell updates/sec

Tabular output not generated.

Title: >US-09-020-716-4
Description: (1:5511) from USG020716 seq
Perfect Score: 5511
N.A. Sequence: 1 TCGCGCGTTTGGTGATGAC . . . TATCAGGAGGCCCTTTGTC 5511
Comp: ACGCGCAAGCCACTACTG . . . ATAGTGTCCGCGAAGCAG

Scoring table:
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htrg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_v1

Database: genbank110
16:gb_bal 17:gb_ba2 18:gb_btg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_v1

Statistics: Mean 12.522; Variance 6.223; scale 2.012

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description Pred. No.
1	2243	40.7	2689 32 Cloning vector pUC-JK, 0.00e+00
2	2241	40.7	4102 32 Cloning vector pBRAY-2, 0.00e+00
3	2242	40.7	4506 32 CVPFCATPRO Cloning vector pCAT-Pr 0.00e+00
4	2242	40.7	4884 32 I70383 Sequence 1 from patent 0.00e+00
5	2242	40.7	4883 22 I82375 Sequence 4 from patent 0.00e+00
6	2242	40.7	4883 22 A18053 DNA for transforming m 0.00e+00
7	2242	40.7	4883 22 A18058 Sequence 4 from patent 0.00e+00
8	2242	40.7	4883 22 CVPFCATPRO Cloning vector pBRAY-1, 0.00e+00
9	2241	40.7	5399 22 I49885 Sequence 1 from patent 0.00e+00
10	2241	40.7	5399 22 A18050 DNA for transforming m 0.00e+00
11	2241	40.7	5399 22 I82373 Sequence 1 from patent 0.00e+00
12	2242	40.7	5620 32 I47737 Sequence 4 from patent 0.00e+00
13	2242	40.7	5620 32 A23324 pVF108 plasmid DNA. 0.00e+00

RESULT	LOCUS	1	AF016541	2689 bp	DNA	circular	SYN	04-SEP-1997
DEFINITION	Cloning vector pUC-JK, complete sequence.							
ACCESSION	AF016541							
NID	92353791							
KEYWORDS	Cloning vector pUC-JK.							
SOURCE	Artificial sequence; cloning vectors.							
ORGANISM	Kim, J.J., Brynan, B.E. and Quackenbush, J.							
REFERENCE	1 (bases 1 to 2689)							
AUTHORS	Kim, J.J., Brynan, B.E. and Quackenbush, J.							
TITLE	pUC-JK cloning vector							
JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 2689)							
AUTHORS	Kim, J.J., Brynan, B.E. and Quackenbush, J.							
TITLE	Direct Submission							
JOURNAL	Submitted (30-JUL-1997) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20878, USA							
FEATURES	Location/Qualifiers							
Source	1..2689							
	/organism="Cloning vector pUC-JK"							
	/db_xref="taxon:65020"							
	/plasmid="pUC-JK"							
	/note="high copy number plasmid derived from pUC19"							
	/complement(146..472)							
	/codon_start=1							
	/transl_table=11							
	/product="beta-galactosidase alpha peptide"							
	/db_xref="PIR:1233792"							
	/translation="MTITSLFVNCTCTPCTGTGNSLAVVLAFFRWENFGVQL NKLAAHFFFAWPAEEAPTPPSCQLRSINGEWFLMRVLLFLGCSHRIWCTLST ICSDA"							
	/complement(395..455)							
	/note="multiple cloning site adapted from pUC19 for BstXI cloning; contains restriction sites for EcoRI, BstXI, KpnI, SmaI, BamHI, PstI, SalI, BstXI and HindIII"							
	/cdo="Region: multiple cloning site"							

ALIGNMENTS

14	2242	40.7	5620 22	AR007513	Sequence 3 from patent	0.00e+00
15	2242	40.7	5642 22	I70384	Sequence 2 from patent	0.00e+00
16	2242	40.7	5647 16	BACNNR	Bacillus stearothermophilus	0.00e+00
17	2242	40.7	6291 12	A60111	Sequence 4 from patent	0.00e+00
18	2242	40.7	6376 22	I33114	Sequence 5 from patent	0.00e+00
19	2242	40.7	6376 22	A23633	Plasmid pVF149.	0.00e+00
20	2242	40.7	6555 22	A8607512	Sequence 2 from patent	0.00e+00
21	2242	40.7	7049 22	A23998	pB4 expression vector	0.00e+00
22	2242	40.7	11490 32	CVU80929	Cloning vector pBACE4.	0.00e+00
23	2242	40.7	16297 32	CVU75991	Cloning vector pBACE4.	0.00e+00
24	2242	40.7	19500 32	CVU75992	Cloning vector pMECA.	0.00e+00
25	2240	40.6	2860 32	AF017063	Bacteriophage phi-X174	0.00e+00
26	2239	40.6	3164 23	BPLVSPR	Cloning vector p21P1.	0.00e+00
27	2240	40.6	3409 32	CVU94951	Cloning vector pUCGM 9	0.00e+00
28	2238	40.6	3563 32	XU704510	Cloning vector pADant	0.00e+00
29	2240	40.6	4392 32	CVU047244	Shuttle vector pCX111.	0.00e+00
30	2238	40.6	4446 32	XU734310	Shuttle vector pDAF2/S	0.00e+00
31	2240	40.6	4630 32	CVU94950	Cloning vector pDAF2.	0.00e+00
32	2238	40.6	4630 32	CVU35316	Cloning vector pX1018G	0.00e+00
33	2238	40.6	5008 32	XU17501	Cloning vector pX1018G	0.00e+00
34	2238	40.6	5154 32	XU17502	Cloning vector pX1018G	0.00e+00
35	2238	40.6	5257 32	XU334320	E.coli/Kluyveromyces 1	0.00e+00
36	2238	40.6	5445 32	AF025392	Cloning vector pUC-GM-	0.00e+00
37	2238	40.6	5862 32	XU334311	E.coli/Kluyveromyces 1	0.00e+00
38	2238	40.6	6461 32	XU334315	E.coli/Kluyveromyces 1	0.00e+00
39	2240	40.6	6687 32	AB012282	Shuttle vector pSVbeta	0.00e+00
40	2238	40.6	6894 32	XU02435	Shuttle vector pSVbeta	0.00e+00
41	2238	40.6	7561 32	XU334313	E.coli/Kluyveromyces 1	0.00e+00
42	2239	40.6	9208 32	AF064057	Expression vector pYZ8	0.00e+00
43	2239	40.6	9211 32	AF064066	Expression vector pYZ4	0.00e+00
44	2239	40.6	9215 32	AF049083	Expression vector pYZ1	0.00e+00
45	2239	40.6	9956 32	AF049084	Expression vector pYZ3	0.00e+00


```

DRYDIALATEDIAEELGGEWADRFVLYGIAAPDSORIAFYRLDEFF"
3485..3490
/Note="hsv thymidine kinase polyA signal"
polyA_signal
3498..3503
/Note="hsv thymidine kinase polyA signal"
polyA_signal
3572..3605
/Note="loxP element"
misc_feature
3674..3802
/Note="multiple cloning site (ClaI, BglII, Pst-EII, NotI)"
misc_feature

BASE COUNT 967 a 1059 c 1085 g 991 t
ORIGIN
Query Match 40.7%; Score 2341; DP 37; Length 4102;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGGAAGGCGCTCGTATACGCTATTTTATAGTTTAAATGTCATGATATAATGTTT 60
Cp 5511 GAGGAAGGCGCTCGTATACGCTATTTTATAGTTTAAATGTCATGATATAATGTTT 5452
Db 61 CTAGAGCTCAGTGGCGCTTTTCGCGAAATGTAAGGGAAGGCGCTATTTGTTTATTTT 120
Cp 5451 CTAGAGCTCAGTGGCGCTTTTCGCGAAATGTAAGGGAAGGCGCTATTTGTTTATTTT 5392
Db 121 TCTAAATACATTCATTAATGATGCTCTCATGAGCAATACCTGATAAATGCTTCAAT 180
Cp 5391 TCTAAATACATTCATTAATGATGCTCTCATGAGCAATACCTGATAAATGCTTCAAT 5332
Db 181 AATATGAAAGGAGAGATGATGATTAATCAACATTCGCTGATGAGGCTTATTCCTTTT 240
Cp 5331 AATATGAAAGGAGAGATGATGATTAATCAACATTCGCTGATGAGGCTTATTCCTTTT 5272
Db 241 TTGGGCGATTTTGGCTTCCTTGTTCCTCACCCAGAAAGGCTGTGTAAGTAAGATG 300
Cp 5271 TTGGGCGATTTTGGCTTCCTTGTTCCTCACCCAGAAAGGCTGTGTAAGTAAGATG 5212
Db 301 CTGAAGATCAGTTCGGTGCAGAGTGGTTTACATCAACATGATCTCAACAGCGGTAGA 360
Cp 5211 CTGAAGATCAGTTCGGTGCAGAGTGGTTTACATCAACATGATCTCAACAGCGGTAGA 5152
Db 361 TCGTTGAGAGTTTTCGCGCGGAGAGGTTTTCATATGATGATGATGATGATGATGATG 420
Cp 5151 TCGTTGAGAGTTTTCGCGCGGAGAGGTTTTCATATGATGATGATGATGATGATGATG 5092
Db 421 TATGTCGCGGATATTCCTGATATGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Cp 5091 TATGTCGCGGATATTCCTGATATGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5032
Db 481 ACTATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Cp 5031 ACTATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4972
Db 541 GCATGACAGTAAGAGAAATATGCGATGCTGCATACCATGATGATGATGATGATGATGATG 600
Cp 4971 GCATGACAGTAAGAGAAATATGCGATGCTGCATACCATGATGATGATGATGATGATGATG 4912
Db 601 ACTATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Cp 4911 ACTATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4852
Db 661 GGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Cp 4851 GGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4792
Db 721 ACAGAGGTGACACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Cp 4791 ACAGAGGTGACACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4732
Db 781 GCGAATCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 840
Cp 4731 GCGAATCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 4672
Db 841 TTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900

```

```

LOCUS 2 CUG63120 4102 bp DNA circular SYN 12-DEC-1996
DEFINITION Cloning vector pRAY-2, complete sequence.
ACCESSION U63120
NID g1724066
KEYWORDS Cloning vector pRAY-2.
SOURCE Cloning vector pRAY-2.
ORGANISM artificial sequence; cloning vectors.
REFERENCE 1 (bases 1 to 4102)
Storck, T., Krueh, U., Kolhekar, R., Sprengel, R. and Seeburg, P.H.
Rapid construction in yeast of complex targeting vectors for gene
manipulation in the mouse
Nucleic Acids Res. 24 (22): 4594-4596 (1996)
JOURNAL 97105907
MEDLINE
AUTHORS Storck, T., Krueh, U., Kolhekar, R., Sprengel, R. and Seeburg, P.H.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1996) Center for Molecular Biology Heidelberg
(2MBH), University of Heidelberg, Im Neuenheimer Feld 289,
Heidelberg 69120, Germany
FEATURES
Source
1. 4102
/organism="Cloning vector pRAY-2"
/plasmid="pRAY-2"
/db_xref="taxon:50337"
/Note="for Rapid Assembly in Yeast"
Join(1..2239,3703..4102)
/Note="plasmid backbone from pUC19"
2240..2270
/Note="multiple cloning site (SfiI, SalI, HindIII, XhoI)"
2301..2334
/Note="loxP element"
2455..2654
/Note="hsv thymidine kinase promoter"
2655..2670
/Note="sequence modified to support efficient translation
in prokaryotes and eukaryotes (Ahmed, A.-Q. and
Mensa-Wilmoth, K. (1996) Nucleic Acids. Res. 24, 1173-1174)"
2671..3474
/codon_start=1
/product="neomycin phosphotransferase"
/db_xref="PID:g1724067"
/translation="NGSAIEQDGLHAGSPAAWVERFGYDWAQQTGCCDAANVFPLSA
QGRVFLVFKYTLDSALNEQDEAARLSNLAATGVPVCAALVDVYVTEAGRWLLGVPG
QDLSSHLPAPAKVSIADAMPPLHLPATCPFDHQAHRERIERATMEAGLVQDDQ
LDEHQGLAPAELEFARKLARPMDGEDLVVTHGDACLPLINVENGREPSGFDGRLGVA

```


[illegible]

RESULT		5			
LOCUS	182376		4883 bp	DNA	PAT
DEFINITION	Sequence 4 from patent US 5712135.				
ACCESSION	182376				
NID	G3210573				
KEYWORDS	UNKNOWN.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				


```

NID 42472108
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4883)
AUTHORS D'Halluin, K. and Gubel, E.
TITLE Process for transforming monocotyledonous plants
JOURNAL Patent, US 5,641,664 A 4 24 JUN 1997.
FEATURES
    Location/Qualifiers
        1..4883
        /organism="unknown"
BASE COUNT 1401 a 1256 c 1201 g 1125 t
ORIGIN
    2642 5'CAAGCTTGGGTAATCATGCTGATAGCTGTTGCTGTGGTAAATGTTATCGGCTCACCA 2701
    3270 5'CAAGCTTGAAGTAATCATGCTGATAGCTGTTGCTGTGGTAAATGTTATCGGCTCACCA 3329
    3762 5'ATTCACACACACATAGTACGCGGAAATATAAAGCTGAAGGCTGGGTGGCTTAAGTAGTG 2761
    4430 5'ATTCACACACACATAGTACGCGGAAATATAAAGCTGAAGGCTGGGTGGCTTAAGTAGTG 3489
    2742 5'ATTAACACACATTAATTCGTTGGCTCAGCTGCGGCTTTCGAGCTGGGAAACCTGTGCG 2821
    4430 5'ATTAACACACATTAATTCGTTGGCTCAGCTGCGGCTTTCGAGCTGGGAAACCTGTGCG 3449
    2822 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2881
    4450 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3509
    2882 5'GTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCT 2941
    4510 5'GTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCT 3569
    2942 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3001
    3570 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3629
    3002 5'ATCATGTGATGAAAAGGAGAGAAAAGGCGAGAAACGCTGAGGAGGCGGCTTTCGCTGGCG 3061
    4630 5'AAATGTTAAATAAAAGGCGAAATAAAAGGCGAGAAACGCTGAGGAGGCGGCTTTCGCTGGCG 3689
    3062 5'GTTTCCCATAGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3121
    4650 5'GTTTCCCATAGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3749
    3122 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3181
    3750 5'GTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3809
    3182 5'GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3241
    4810 5'GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3869
    3242 5'AGGCTGGGCTTTCCTCAATAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 3301
    3870 5'AGGCTGGGCTTTCCTCAATAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 3929
    3302 5'GCAAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3361
    4830 5'GCAAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3989
    3362 5'AAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3421
    4890 5'AAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4049
    3422 5'GTAACAGGATTAATGAGAGGATTAATGAGAGGATTAATGAGAGGATTAATGAGAGGATTAATGAGAGG 3481

```

```

Db 4562 CGTGGACCAACATGATTCAGCATCTTTTACTTTCACGACGCTTTCGGTGACGAAA 4621
|||||
QY 5190 CGTGGACCAACATGATTCAGCATCTTTTACTTTCACGACGCTTTCGGTGACGAAA 5249
|||||
Db 4624 ACAGCAAAATTAATATCGGCAAAAGGAATAAGGGCGGACGCGGAATGTTCAATATCTC 4681
|||||
QY 5250 ACAGCAAAATTAATATCGGCAAAAGGAATAAGGGCGGACGCGGAATGTTCAATATCTC 5309
|||||
Db 4682 ATACTCTTCTTTTTCATATATTATTAAGCAATTTATCAGGCTTATGCTCATGACGCGA 4741
|||||
QY 5310 ATACTCTTCTTTTTCATATATTATTAAGCAATTTATCAGGCTTATGCTCATGACGCGA 5369
|||||
Db 4742 TACATATTTGAATGATTTAGAAAATAAACAATAGAGGCTTCGCGCAATTTCCCGCA 4801
|||||
QY 5370 TACATATTTGAATGATTTAGAAAATAAACAATAGAGGCTTCGCGCAATTTCCCGCA 5429
|||||
Db 4802 AAGTGCACCTGACGCTTAAGAACCATTTATTCATGACATTAACCTATAAAATAGG 4861
|||||
QY 5430 AAGTGCACCTGACGCTTAAGAACCATTTATTCATGACATTAACCTATAAAATAGG 5489
|||||
Db 4862 CCAATACAGAGGCTTTTGGTC 4893
|||||
QY 5490 CCAATACAGAGGCTTTTGGTC 5511
|||||

RESULT 8 CU063018 5225 bp DNA circular SYN 12-DEC-1996
LOCUS Cloning vector pRAY-1, complete sequence.
DEFINITION U63018
ACCESSION 63724062
XREF 63724062
KEYWORDS Cloning vector pRAY-1
SOURCE Cloning vector pRAY-1
ORGANISM artificial sequence; cloning vectors.
REFERENCE 1 (bases 1 to 5225)
AUTHORS Sterckx, J., Kruth, U., Kolhekar, R., Sprengel, R. and Seeburg, P.H.
TITLE Rapid construction in yeast of complex targeting vectors for gene
manipulation in the mouse
JOURNAL Nucleic Acids Res. 24 (22), 4594-4596 (1996)
MEDLINE 47105407
REFERENCE 2 (bases 1 to 5225)
AUTHORS Sterckx, J., Kruth, U., Kolhekar, R., Sprengel, R. and Seeburg, P.H.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1996) Center for Molecular Biology Heidelberg
(2MBH), University of Heidelberg, Im Neuenheimer Feld 282,
Heidelberg 69120, Germany
FEATURES
source
1..5225
/organism="Cloning vector pRAY-1"
/plasmid="pRAY-1"
/notes="for Rapid Assembly in Yeast"
/db_xref="taxon:5047"
/lab_host="Saccharomyces cerevisiae"
join(1..2239,4826..5225)
/notes="plasmid backbone derived from pUC19"
201..4061
/function="ampicillin resistance"
/codon_start=1
/product="beta-lactamase"
/db_xref="PID:q1724063"
/translation="MSTQHPRVALIPFFAAFCPLVFAHPETLVKVKDAEDQLGARVGY
IELDNKKILLESPPFPFMSTFKVLICGLAVLSRIDAGSEGLGRIRHYSONDLVE
VSPVTEKIDMTQIVRELCSAAITMSDNTAANLLLTITIGKPKELAFLEHMGDHWRL
DRWPEPELNEATPDEPTTMPVMAWTLTPKILTGELLTLASFOQLTDWEADKVAQPL
LRSLAPGAWFVADKSGAGERSGIITAAALGPDGPKSRIVIVITGSGATMDERNQIA
EIGASLIRKW"
2240..2270
/notes="multiple cloning site (SfiI, SalI, HindIII, XhoI)"
2301..2334
/notes="encodes loxP element"
2456..2661
/notes="HSV thymidine kinase"

```

```

CDS
2662..3455
/codon_start=1
/product="neomycin phosphotransferase"
/db_xref="PID:q1724064"
/translation="MGSALDUGGLHAGSFAANYERLFVYEWKALITDSEDAATFSLA
QGRPLVKTDLSCALNELODEARLSWLAITGVCAPAAVDVVTGAGRWLLSEVPG
QGLSSHLAFAEKVSINADAMRLRLDRLATFFEHCAKHFLEPAFTMEALVQDDE
LDREHGLAPAELEFARLKARMPDGEDLVTHGACLPNFWENGSPFSQFIDGGRJGVA
DRYQDIALATRLDIAELGGKADREFLVLCIAPDSDQRIATFYPLLDFF"
polyA_signal
3476..3481
/notes="putative; HSV thymidine kinase"
polyA_signal
3489..3494
/notes="putative; HSV thymidine kinase"
gene
3766..4569
/genes="URA3"
CDS
3766..4569
/genes="URA3"
/notes="yeast URA3 selection marker"
/codon_start=1
/product="orotidine-5'-phosphate decarboxylase"
/db_xref="PID:q1724065"
/translation="MSKATYKERAATHPSVAAKLFNIMHEKOTNLCASLDVRTTKEL
LELVALSPKICLLKTHVGLTDFSMGCTVFPKALSAPVNFLLPFRFACIGNTVK
LQTSQVIRIAEWADITNAHGVSPGIYSSGLKQAAEVTKPEGLIMLALSPKSLLS
TGEYTKGIVDIAKSDKDEIVGFIADROMXGRKDEYEWLIMHP-SVGLDEKSEALQQVR
TVDDVVSIGSDIIIVGRGLFAGDEDAKVEGEYPYKAGWEAYFLRRCGQCN"
misc_feature
4695..4728
/notes="encodes loxP element"
misc_feature
4826..5225
/notes="multiple cloning site (ClaI, BglII, BstEII, NotI)"
BASE COUNT 1323 a 1251 c 1350 g 1301 t
ORIGIN
Query Match 40.7%, Score 2241, DB 12, Length 5225,
Best Local Similarity 100.0%, Pred. No. 0.00e+00;
Matches 2242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 GACGAAAGGGCTCGTATAGCGCTATTTTATAGCTTAATGTCATGATAAATGGTIT 60
|||||
Cp 5511 GACGAAAGGGCTCGTATAGCGCTATTTTATAGCTTAATGTCATGATAAATGGTIT 5452
|||||
Db 61 GTTAGAGCTGAGTGGGCACTTTTCGGGAAATGTGCGGCAACCCCTATTGTTTATTT 120
|||||
Cp 5451 GTTAGAGCTGAGTGGGCACTTTTCGGGAAATGTGCGGCAACCCCTATTGTTTATTT 5492
|||||
Db 121 TCTAATACATTCAAATATGATCGCTCATGAGACATAAACCCTGATTAATGCTCAAT 180
|||||
Cp 5391 TCTAATACATTCAAATATGATCGCTCATGAGACATAAACCCTGATTAATGCTCAAT 5332
|||||
Db 181 AATATTGAAAAGGAGAGATGATGATTAACAAITTCGCGTGGGCTTATTCGCTTTT 240
|||||
Cp 5331 AATATTGAAAAGGAGAGATGATGATTAACAAITTCGCGTGGGCTTATTCGCTTTT 5272
|||||
Db 241 TTCGGCATTTTGCCTTCCTGTTTTCGCTACGAGAAACGCTGCTGAAAGCAATG 300
|||||
Cp 5271 TTCGGCATTTTGCCTTCCTGTTTTCGCTACGAGAAACGCTGCTGAAAGCAATG 5212
|||||
Db 301 CTGAGATCAGTGGGTGCGACAGTGGGTATCATGCAACGAGGAGTCAACAGCGGTAA 360
|||||
Cp 5211 CTGAGATCAGTGGGTGCGACAGTGGGTATCATGCAACGAGGAGTCAACAGCGGTAA 5152
|||||
Db 361 TCCTTGAGAGTTTCGCCCGGCAAGACGTTTTCCTAATCATGAGCACTTTTAAAGTTC 420
|||||
Cp 5151 TCCTTGAGAGTTTCGCCCGGCAAGACGTTTTCCTAATCATGAGCACTTTTAAAGTTC 5062
|||||
Db 421 TAGTGGCGGGTATTAATCGGCTATGATGAGTGGTGGTCAAGCAATGAGTGGTGG 480
|||||
Cp 5091 TAGTGGCGGGTATTAATCGGCTATGAGTGGTGGTGGTCAAGCAATGAGTGGTGG 5032
|||||
Db 481 ACTATTCTCAGATGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
|||||
Cp 5031 ACTATTCTCAGATGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4972
|||||

```


RESULT 10 A18756 5494 bp UNA PAT 65-APP-1944
 DEFINITION DNA for transforming monocotyledonous plants seq ID No.11.
 ACCESSION A18756
 NID 4513154
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 unclassified.
 REFERENCE 1 (bases 1 to 5394)
 AUTHORS
 JORDAN L
 FEATURES
 source 1..5394 "Artificial sequence"
 misc_feature 1..451
 /note="puc18 derived sequence"
 promoter 452..1284
 /note="15S3 promoter sequence derived from cauliflower mosaic virus isolate cabB31"
 gene 1285..2100
 /gene="neomycin phosphotransferase gene"
 /note="neomycin phosphotransferase gene"
 /codon_start=1
 /cdd_xref="P10_4513155"
 /translation="MGLKAWLWLPGLGLHAGSSAAWVSELSYLNWAAQLUGSDAAVF
 RLSDGRPVLFVRIDLSALNELDRAAPRI SWIATTCVDPAAVIVVTFAPRDMWLLIG
 EVQDDIL SRIHAPKPVSMADAMPRRI HTI DPATC PFEDQAKRIETAPRTMEAGLV
 IQDDLDPEHGLIAPAEI FARI KAPMPFSEI VVHRCATLPNIMVFNIPFSFIUNGR
 LGVAPRYQDIALATROIAPFELGGEWALRFLVLYGIAAPDSQRIATFVRLDDEF"
 misc_feature 3161..5399
 /note="puc18 derived sequence"
 BASE COUNT 1359 a 1378 g 1369 t 1244 c
 ORIGIN
 Query Match: 40.7% Score 2241; DB 22;
 Best Local Similarity 100.0% E-val No. 5.0e-005
 Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 3159 CAAGCTTGCTGAATGATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3218
 QY 4271 CAAGCTTGCTGAATGATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3330
 DB 4219 TTTGTAACAAATATGATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3278
 QY 4441 TTTGTAACAAATATGATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3390
 DB 4279 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3338
 QY 4391 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3450
 DB 4339 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3398
 QY 4451 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3510
 DB 4399 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3458
 QY 4511 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3570
 DB 4459 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3518
 QY 4571 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3630
 DB 4519 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3578
 QY 4631 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3690
 DB 4579 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3638
 QY 4691 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 3750

DB 3639 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4098
 QY 3751 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4130
 DB 3699 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4158
 QY 3811 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4230
 DB 3759 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4258
 QY 3871 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4330
 DB 3819 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4358
 QY 3931 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4430
 DB 3879 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4458
 QY 3991 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4530
 DB 3939 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4558
 QY 4051 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4630
 DB 3999 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4658
 QY 4111 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4730
 DB 4059 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4758
 QY 4171 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4830
 DB 4119 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4858
 QY 4231 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4930
 DB 4179 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 4958
 QY 4291 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5030
 DB 4239 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5058
 QY 4351 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5130
 DB 4299 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5158
 QY 4411 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5230
 DB 4359 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5258
 QY 4471 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5330
 DB 4419 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5358
 QY 4531 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5430
 DB 4479 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5458
 QY 4591 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5530
 DB 4539 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5558
 QY 4651 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5630
 DB 4599 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5658
 QY 4711 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5730
 DB 4659 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5758
 QY 4771 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5830
 DB 4719 GCTAAGTCAATTAATGCTGATAGCTGTTCTGTGTGTGTAATGTTATCGGCTACAA 5858

QY	4831	CAAGCGAGTACATGATCCCGCATGTTGTCACAAAGCGGTTAGCTCCTTCGGTCTTC	4890	3271	CAAGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGMAAATTGTTATTCGGTCACAA	3330
DB	4779	CGATCGTGTGAGAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGC	4838	3219	TTCCACAAACATAGAGCGGAAAGCATAAAGTGTAAAGCGTGGGTTGCTTAATGAGTGA	3278
QY	4891	CGATCGTGTGAGAGTAAGTTGGCCGAGTGTATCACTCATGTTTATGGCAGCACTGC	4950	3331	TTCCACAAACATAGAGCGGAAAGCATAAAGTGTAAAGCGTGGGTTGCTTAATGAGTGA	3390
DB	4839	ATAATCTCTTACTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC	4898	3279	GTATCTACATTAATTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3338
QY	4951	ATAATCTCTTACTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC	5010	3391	GTATCTACATTAATTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3450
DB	4899	CAAGTCACTTCTGAGTAAGTGTATGCGCGAGCGAGCTTCTGCTGCGCGGCTCAATAC	4958	3339	GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3398
QY	5011	CAAGTCACTTCTGAGTAAGTGTATGCGCGAGCGAGCTTCTGCTGCGCGGCTCAATAC	5070	3451	GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3510
DB	4959	GGGATAATACCGCGGACATAGCAGAACTTTAAAGTGTCTCATCATTTGGGAAAGCTTCT	5018	3399	CTTCGCTTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3458
QY	5071	GGGATAATACCGCGGACATAGCAGAACTTTAAAGTGTCTCATCATTTGGGAAAGCTTCT	5130	3511	CTTCGCTTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3570
DB	5019	CGGCGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5078	3459	CAGCTCACTCAAAAGGCGGTAAATAGGTTATCCACAGAAATCAGGCGGTAAATCAGGAAAGA	3518
QY	5131	CGGCGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5190	3571	CAGCTCACTCAAAAGGCGGTAAATAGGTTATCCACAGAAATCAGGCGGTAAATCAGGAAAGA	3630
DB	5079	GTCAGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5138	3519	ACATGTGAGCAAAAGGCGGTAAATAGGTTATCCACAGAAATCAGGCGGTAAATCAGGAAAGA	3578
QY	5191	GTCAGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5250	3631	ACATGTGAGCAAAAGGCGGTAAATAGGTTATCCACAGAAATCAGGCGGTAAATCAGGAAAGA	3690
DB	5139	CAGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5198	3579	TTTTCATAGGCTCCGCGCGCTGAGGAGTATCAGGAAATCAGGCGGTAAATCAGGAAAGA	3638
QY	5251	CAGGAAAGCTCTCAAGGATCTTACCGCTGTTTACGATTCAGTTCGATGTAACCCACTC	5310	3691	TTTTCATAGGCTCCGCGCGCTGAGGAGTATCAGGAAATCAGGCGGTAAATCAGGAAAGA	3750
DB	5199	TACTTCTCTTTTCAATATTTAAGGATTTATCAGGCTTATGTCATGAGCGGAT	5258	3639	GGGAAAGCTCTCAAGGATTTAAGATACAGGCGGTTCCTCCCTGGAAGCTCTGCTGCTGCT	3698
QY	5311	TACTTCTCTTTTCAATATTTAAGGATTTATCAGGCTTATGTCATGAGCGGAT	5370	3751	GGGAAAGCTCTCAAGGATTTAAGATACAGGCGGTTCCTCCCTGGAAGCTCTGCTGCTGCT	3810
DB	5259	ACATATTTGATGTTTACGAAATTAACAAATAGGCTTCCGCGGACATTTCCCGGAA	5318	3699	GTCTCTCTGTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT	3758
QY	5371	ACATATTTGATGTTTACGAAATTAACAAATAGGCTTCCGCGGACATTTCCCGGAA	5430	3811	GTCTCTCTGTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT	3870
DB	5319	AAGTCTGATGTTTACGAAATTAACAAATAGGCTTCCGCGGACATTTCCCGGAA	5378	3759	GGGTGGCGCTTCTCAATGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATG	3818
QY	5431	AAGTCTGATGTTTACGAAATTAACAAATAGGCTTCCGCGGACATTTCCCGGAA	5490	3871	GGGTGGCGCTTCTCAATGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATG	3930
CL	5379	STATCAGAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT		3819	CCAACTGGGCTGTTGTCAGGAGCGGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT	3878
QY	5491	STATCAGAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT		3931	CCAACTGGGCTGTTGTCAGGAGCGGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTT	3990
RESULT	11	182373	5399 bp	DNA	PAT	20-MAR-1998
LOCUS	182373	Sequence 1 from patent US 5712135.				
DEFINITION	182373					
ACCESSION	g3210670					
NID	Unknown.					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1. (bases 1 to 5399)					
AUTHORS	D'Halluin, K. and Gobel, E.					
TITLE	Process for transforming monocotyledonous plants					
JOURNAL	Patent: US 5712135-A 1 27-JAN-1998;					
FEATURES	Location/Qualifiers					
source	1..5399					
BASE COUNT	1359 a 1378 c 1369 g 1293 t					
ORIGIN						
Query Match	40.7%	Score	2241.	DB	22	Length 5399:
Best Local Similarity	100.0%	Pred. No.	0.00e+00;			
Matches	2241;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
DB	3159	CAAGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGMAAATTGTTATTCGGTCACAA	3218			

Db	4299	AATCAATCTAAAGTATATAGTAAGTAAGTCTGTGAGAGTACCAATGCTTATACAGTG	4358
Qy	4411	AATCAATCTAAAGTATATAGTAAGTAAGTCTGTGAGAGTACCAATGCTTATACAGTG	4470
Db	4358	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4418
Qy	4471	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4530
Db	4419	TGATATAACTATATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	4479
Qy	4531	TGATATAACTATATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	4590
Db	4479	SATACCAACTATATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	4538
Qy	4591	SATACCAACTATATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	4650
Db	4538	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4598
Qy	4651	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4710
Db	4598	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4658
Qy	4711	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4770
Db	4658	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4718
Qy	4771	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4830
Db	4718	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4778
Qy	4831	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4890
Db	4778	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4838
Qy	4891	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4950
Db	4838	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4898
Qy	4951	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5010
Db	4898	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	4958
Qy	5011	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5070
Db	4958	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5018
Qy	5071	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5130
Db	5018	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5078
Qy	5131	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5190
Db	5078	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5138
Qy	5191	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5250
Db	5138	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5198
Qy	5251	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5310
Db	5198	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5258
Qy	5311	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5370
Db	5258	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5318
Qy	5371	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5430
Db	5318	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5378
Qy	5431	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	5490
Db	5378	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	
Qy	5491	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	
Db	5490	AGCAATCTACAGGATCTGTGATTTCTGTATGATATAGTATGCTGCTGCTGCTGCTG	
Qy			
Db	5379	GTATCAGGAGGCGCTTTTCGTC	5399
Qy	5491	GTATCAGGAGGCGCTTTTCGTC	5511
RESULT	12		
LOCUS	147737	5620 bp	DNA
DEFINITION	Sequence 9 from patent US 5639448.		29 JUL 1999
ACCESSION	147737		
KEYWORDS	g2471702		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 5620)		
AUTHORS	Michiels, F., Morioke, S., Subeljnick, T. and Komari, T.		
TITLE	Stamen-specific promoters from rice		
JOURNAL	Patent: US 5639448-A 9 17-JUN-1997		
FEATURES	Location/Qualifiers		
SOURCE	1..5620		
BASE COUNT	1498 a 1359 c 1429 g 1430 t		4 others
ORIGIN	/organism="unknown"		
Query Match:	40.7%	Score: 2242	DB 22; Length: 5620;
Best Local Similarity	100.0%	Pred. No. 0.00e+00;	
Matches	2242;	Conservative 0;	Mismatches 0;
Models	0;	Gaps	0;
Db	3379	GAAGCTTGGGATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3438
Qy	3276	GAAGCTTGGGATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3439
Db	3439	ATTCACACACATACAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3498
Qy	3330	ATTCACACACATACAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3489
Db	3499	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3558
Qy	3390	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3449
Db	3559	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3618
Qy	3450	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3609
Db	3619	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3678
Qy	3510	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3669
Db	3679	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3738
Qy	3570	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3729
Db	3739	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3798
Qy	3630	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3809
Db	3799	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3858
Qy	3690	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3849
Db	3859	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3918
Qy	3750	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3909
Db	3919	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3978
Qy	3810	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	3969
Db	3979	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	4038
Qy	3870	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	4029
Db	4039	AGCAATCTACATTAATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGATCAAGGAT	4098

```
QY 3930 ACCAAGTCATTCTGAGAAATAGTGTATGCGGCGACCGAGTTGTCCTTTCGCCGGCGTCAATA 5069
DB 5179 CCGGATAATAATGCGGCGACACA:AGCAGAACTTTAAAGATGTTTAAAGATGCGGCGTCTT 5238
QY 5070 CCGGATAATAATGCGGCGACACA:AGCAGAACTTTAAAGATGTTTAAAGATGCGGCGTCTT 5129
DB 5239 TCGGCGGCGGAAACTGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGG 5298
QY 5130 TCGGCGGCGGAAACTGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGG 5199
DB 5299 CTTGCAACCAACTGATCTTCAGCATCTTTTACCTTACAGAGAGAGAGAGAGAGAGAGAGAG 5358
QY 5190 CTTGCAACCAACTGATCTTCAGCATCTTTTACCTTACAGAGAGAGAGAGAGAGAGAGAGAG 5249
DB 5359 ACAGGAGGCGGAAACTGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGG 5418
QY 5250 ACAGGAGGCGGAAACTGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGGAGATGTCAGG 5309
DB 5419 ATACTCTCTCTTTTCAATATTTTGAAGCAATTTTGAAGCAATTTTGAAGCAATTTTGAAGCA 5478
QY 5310 ATACTCTCTCTTTTCAATATTTTGAAGCAATTTTGAAGCAATTTTGAAGCAATTTTGAAGCA 5369
DB 5479 TACATTTTGAATGTTTATTTAGAAAATTAACCAATTAACCAATTAACCAATTAACCAATTAAC 5538
QY 5370 TACATTTTGAATGTTTATTTAGAAAATTAACCAATTAACCAATTAACCAATTAACCAATTAAC 5429
DB 5539 AAGGCGGCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 5598
QY 5430 AAGTGGCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 5489
DB 5599 CGTATCGAGGCGGCGTCTTCTCT 5620
QY 5490 CCTATCGAGGCGGCGTCTTCTCT 5511

RESULT 13 A2334 5620 bp LNA PAT 01-APR-1995
LOCUS PVE108 plasmid DNA.
DEFINITION A2334
ACCESSION A2334
KEYWORDS g904321
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
REFERENCE 1 (bases 1 to 5620)
AUTHORS STAMEN-SPECIFIC PROMOTERS FROM RICE
TITLE Patent. WO 9213956-A 9 20-AUG-1992.
JOURNAL Escherichia.
FEATURES
    source
        1..5620
        /organism="Escherichia coli"
        /plasmid="Plasmid pVE108"
        /db_xref="taxon:562"
    misc_feature
        1..395
        /note="pUC18 derived sequence"
    misc_feature
        396..802
        /note="Agrobacterium T-DNA derived sequence: contains
        polyA site from nopaline gene"
    promoter
        1139..1683
        /note="tapetum-specific: Nicotiana tabacum derived
        sequence"
    promoter
        1684..2516
        /note="35S3: Callitower mosaic virus isolate CabE8-31
        derived sequence"
    gene
        2517..3068
        /gene="bar"
    CDS
        2517..3068
        /codon_start=1
        /transl_table=1
        /product="phosphotriester transferase"
        /db_xref="PIR:9904322"
```


QY 4290 TTGATCTTTTCTAGCGGGGICGAGGCGICAGTGGAGCAAAAACCTCACGTTAAGGATTTG 4349
DB 4481 GTCATGAGATTAICAAAAAGATGCTTACCTAGATGCTTTTAAATTTAAAATGAAGTTT 4540
QY 4350 GTCATGAGATTAICAAAAAGATGCTTACCTAGATGCTTTTAAATTTAAAATGAAGTTT 4409
DB 4541 AAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTAAACCAATGTTAAACAGT 4600
QY 4410 AAATCAATCTAAAGTATATATAGTAAACCTTGGTCTGACAGTAAACCAATGTTAAACAGT 4469
DB 4501 SAGGACCTATATACAGGAGTGTGCTGATGCTTATGCTGATGCTGATGCTGATGCTGCT 4660
QY 4470 SAGGACCTATATACAGGAGTGTGCTGATGCTTATGCTGATGCTGATGCTGATGCTGCT 4529
DB 4661 GTGTAGATAAGTATGATGAGGAGGAGTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4720
QY 4530 GTGTAGATAAGTATGATGAGGAGGAGTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4589
DB 4721 CGAGACCCAGCTCACCGGCTGCTGATTTATCAGCAATATAACAGCCAGCCGGAAGGCC 4780
QY 4590 CGAGACCCAGCTCACCGGCTGCTGATTTATCAGCAATATAACAGCCAGCCGGAAGGCC 4649
DB 4781 GAGCGCAAAAGTGCTGCTGCAACCTTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4840
QY 4650 GAGCGCAAAAGTGCTGCTGCAACCTTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4709
DB 4841 SAAGTATAGTAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4900
QY 4710 SAAGTATAGTAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4769
DB 4901 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4960
QY 4770 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4829
DB 4961 TCAAGCGGAGTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5020
QY 4830 TCAAGCGGAGTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4889
DB 5021 CGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5080
QY 4890 CGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4949
DB 5081 CATATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5140
QY 4950 CATATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5009
DB 5141 ACTAAGTATATATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5200
QY 5010 ACTAAGTATATATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5069
DB 5201 CGGATATATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5260
QY 5070 CGGATATATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5129
DB 5261 TCGGCGGCAAAAGTCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5320
QY 5130 TCGGCGGCAAAAGTCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5189
DB 5321 CGTGACCCCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5380
QY 5190 CGTGACCCCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5249
DB 5381 ACAGAGAGCAAAATGAGGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGG 5440
QY 5250 ACAGAGAGCAAAATGAGGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGG 5309
DB 5441 ATACTCTTCTTTTCAATATATATGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGGAGGAG 5500
QY 5310 ATACTCTTCTTTTCAATATATATGAGGAGTATAGGAGGAGTATAGGAGGAGTATAGGAGGAG 5369
DB 5501 TACATATTTGAATGATTTAGAAAAATATAACAAATAGGAGGAGTATAGGAGGAGTATAGGAGGAG 5560
QY 5370 TACATATTTGAATGATTTAGAAAAATATAACAAATAGGAGGAGTATAGGAGGAGTATAGGAGGAG 5429

DB 5561 AAAGTCCACCTGACGCTTAAGAAACCATTTATCATGACATTAACTATATAAATAGG 5620
QY 5430 AAAGTCCACCTGACGCTTAAGAAACCATTTATCATGACATTAACTATATAAATAGG 5489
DB 5621 CGTATCAGAGGCGCTTTTGGTC 5642
QY 5490 CGTATCAGAGGCGCTTTTGGTC 5511

Search completed: Mon Mar 15 04:49:13 1999
Job time : 11984 secs.

[REDACTED]

(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, UK
Distribution rights by Oxford Molecular Ltd

MPsrch_nnn n.d. - n.d. database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 07:51:39 1999 Master time 776.99 seconds
964.996 Million cell updates/sec
Tabular output not generated
Title: >US-09-020-716-4
Description: (11-5511) from US09020716 seq
Percent Score: 5511
N.A. Sequence: 1 TGGGCGGTTTGGTTATGAC TATATGCTCGCGGAAGCAG
Comp: AGCGGCGAAGGCACTACTG TATATGCTCGCGGAAGCAG

Scoring table: TABLE default

Gap 5

Nmatch STD : Dbase 0: Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 10.426; Variance 0.246; scale 1.669

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description	Pred. No.
1	2242	40.7	4453	32	Human cyclin D1-human	0.00e+00
2	2241	40.7	4454	30	Plasmid pGEM1.8 cont	0.00e+00
3	2242	40.7	4510	22	Human cyclin D1-human	0.00e+00
4	2242	40.7	4883	8	pOE110, plasmid DNA r	0.00e+00
5	2242	40.7	4983	4	pOE110, plasmid DNA r	0.00e+00
6	2241	40.7	5299	4	pOE110, plasmid DNA r	0.00e+00
7	2242	40.7	5299	4	pOE110, plasmid DNA r	0.00e+00
8	2242	40.7	5620	9	Plasmid pVE108 used i	0.00e+00
9	2242	40.7	5620	9	Plasmid pVE108 used i	0.00e+00
10	2241	40.7	5555	9	Plasmid pVE144 used i	0.00e+00
11	2242	40.7	7053	9	Sequence of pTE4.	0.00e+00
12	2240	40.6	3851	12	Plasmid pSBC-cyt/c DN	0.00e+00
13	2239	40.6	4045	13	Plasmid pND211 contai	0.00e+00

14	2235	40.6	4118	35	T69189	Construct pGEM-htr(Ka	0.00e+00
15	2235	40.6	4118	35	T69188	Construct pGEM-htr co	0.00e+00
16	2235	40.6	487347	14	Plasmid pENV1	Plasmid pENV1	0.00e+00
17	2235	40.6	4621	32	T69362	Human cyclin D1-human	0.00e+00
18	2239	40.6	4824	23	T69376	Shortened c1 gene in	0.00e+00
19	2234	40.6	4295	29	V12347	Plasmid pWP3.12-vec	0.00e+00
20	2238	40.6	8710	17	O89650	pScl1 FIPV E1.	0.00e+00
21	2238	40.6	9020	17	T07452	pScl1 FIPV E1.	0.00e+00
22	2235	40.6	10930	20	O81336	Plasmid pM16-1.	0.00e+00
23	2235	40.6	10930	20	O81335	Plasmid pM16.	0.00e+00
24	2231	40.5	4283	37	T86610	Epidermal growth fact	0.00e+00
25	2231	40.5	4283	35	T86449	DNA encoding hGH. BGF	0.00e+00
26	2232	40.5	4713	39	V12067	Murine IAD alpha chai	0.00e+00
27	2232	40.5	4734	39	V12068	Murine IAD beta chain	0.00e+00
28	2232	40.5	5294	6	O34111	Plasmid pK3.	0.00e+00
29	2230	40.5	6010	2	O11998	Human plasmidogen wit	0.00e+00
30	2234	40.5	6360	23	T35109	Plasmid pUT19-SV3.	0.00e+00
31	2233	40.5	6458	2	N40096	Sequence of plasmid p	0.00e+00
32	2234	40.5	7556	7	O42160	Plasmid pFS012 contg	0.00e+00
33	2234	40.5	7639	7	O42159	Plasmid pFS034 contg.	0.00e+00
34	2226	40.4	2686	30	T51831	pOE18	0.00e+00
35	2226	40.4	4163	5	O32349	Template clone Fab N9	0.00e+00
36	2225	40.4	4603	1	O04010	Plasmid pSUN387	0.00e+00
37	2228	40.4	4910	22	T15387	pREP76/52 hybrid aden	0.00e+00
38	2228	40.4	4950	38	V03801	Patoviral vector con	0.00e+00
39	2228	40.4	5033	12	O74566	pEHI14.1 vector contg	0.00e+00
40	2228	40.4	5376	4	O27951	pVE149.	0.00e+00
41	2224	40.4	7635	31	T71323	Hybrid activation vec	0.00e+00
42	2224	40.4	9534	30	T62072	Vector pR324E2-emb3d	0.00e+00
43	2228	40.4	9632	38	T97898	Baculovirus transfor	0.00e+00
44	2226	40.4	19307	20	T27558	Shuttle vector pADBel	0.00e+00
45	2222	40.3	2755	5	O30566	Vector pNN03	0.00e+00

ALIGNMENTS

RESULT 1
ID T69890 standard: DNA: 4453 BP.

AC T69890:
DI 19-SEP-1997 (first entry)
DE Human cyclin D1-human CLK4 gene fusion in plasmid pK485.
KW Cyclin D1; cyclin-dependent protein kinase 4; CLK4; cell cycle; ss.
OS Chimeric Homo sapiens;
CH Chimeric synthetic.
FH Key Location/Qualifiers
FT cds 2560..4248
ET cds /*tag= a
FN WO725345-A1.
PN 17-JUL-1997.
PD 03-JAN-1997: U00140.
PF 03-JAN-1996: US-009629.
PA (ELIL) LILLY & CO ELI.
PI Kovacevic S, Otto KA, Rao RN;
DR WPI: 97-372814/34.
DR P-PSDB: W18572.
PT Fusion protein comprising human cyclin and cyclin dependent protein
PT kinase - useful for research in cell cycle regulation allowing
PT addition of both components simultaneously, giving greater control;
PT over reaction conditions
PS Disclosure: Page 10-15: 57pp; English.
QC A DNA sequence (T69890) in plasmid K485, deposited in E. coli as
CC NPT: B-2149, includes a coding region for a fusion protein
CC (W18572) comprising human cyclin D1 linked to human CLK4, flanked
CC by N-terminal histidine residues, a myc epitope and a thrombin
CC cleavage site, and by a C-terminal streptavidin binding domain
CC to facilitate purification. The fusion protein can be expressed
CC in host cells, pref. using a baculovirus expression system, for use
CC in research into cell cycle regulation.
SQ Sequence 4453 BP, 1058 A, 1172 C, 1212 G, 1011 T;

Query Match 40.7%; Score 2242, DB 32, Length 4453;
Best Local Similarity 95.8%, Pred. No. 0.00e+00;
Matches 2246, Conservative 0, Mismatches 4, Indels 0, Gaps 0

1140 Db 1081 catatatacttaatttaaatctcaatttttaatttaataaataatctaatataaa 1147
1141 Cp 4431 CATATATACCTTAGATTGATTTAAATCTTCAATTTTAAATTTAAATGAGATCAATGAGAA 4438
1142 Db 1141 tecttttttaataatctctataccataaataatctcaatcaatcttttcaatctcaatcaat 1200
1143 Cp 4371 TCTCTTTTGAATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4378
1144 Db 1201 caaaccccaatataaataatataaataatctcttaataatcttttttaataatcaatct 1208
1145 Cp 4311 CAGACCCCGTACGAAAGATTTAAAGCATCTCTCTTCAATGATGATGATGATGATGATGAT 4318
1146 Db 1261 actgctctcaacaacaaacaaac 1268
1147 Cp 4251 GTCTCTTGGCAATCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4258
1148 Db 1321 tacraactctttttccgaagataaactgctttccacacacacacacacacacacacacacac 1328
1149 Cp 4191 TACCAACCTCTTTTCCGAAAGGTAACCTGGATGATGATGATGATGATGATGATGATGATGAT 4198
1150 Db 1381 ttttaataataac 1440
1151 Cp 4131 TTCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4072
1152 Db 1441 tccctctctcaatct 1448
1153 Cp 4071 TGCT 4078
1154 Db 1501 qattcaactcaaaac 1508
1155 Cp 4011 GCTTGGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4018
1156 Db 1561 cctac 1620
1157 Cp 3951 CGTGCACATAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4052
1158 Db 1621 agcatctgaagaaagc 1680
1159 Cp 3891 AGCATGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4052
1160 Db 1681 cgaagctcgaagaaagc 1740
1161 Cp 3831 CGAGGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4072
1162 Db 1741 ataatct 1800
1163 Cp 3771 ATAGTCT 4172
1164 Db 1801 gaggcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1808
1165 Cp 3711 GGGGGGGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4072
1166 Db 1861 actgaccttttgcac 1920
1167 Cp 3651 GCTGGGCTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4052
1168 Db 1921 ttaccgccttttaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 1980
1169 Cp 3591 TTACGCGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4052
1170 Db 1981 caatgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 2040
1171 Cp 3531 CACTAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4072
1172 Db 2041 caatcatttaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 2100
1173 Cp 3471 CATTCATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4122
1174 Db 2101 aagcaatttaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 2160
1175 Cp 3411 AGCAATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4072
1176 Db 2161 cagctcgaatct 2220


```

14 4661 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4720
15 4662 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4721
16 4663 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4722
17 4664 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4723
18 4665 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4724
19 4666 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4725
20 4667 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4726
21 4668 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4727
22 4669 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4728
23 4670 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4729
24 4671 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4730
25 4672 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4731
26 4673 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4732
27 4674 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4733
28 4675 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4734
29 4676 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4735
30 4677 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4736
31 4678 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4737
32 4679 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4738
33 4680 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4739
34 4681 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4740
35 4682 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4741
36 4683 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4742
37 4684 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4743
38 4685 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4744
39 4686 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4745
40 4687 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4746
41 4688 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4747
42 4689 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4748
43 4690 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4749
44 4691 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4750
45 4692 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4751
46 4693 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4752
47 4694 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4753
48 4695 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4754
49 4696 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4755
50 4697 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4756
51 4698 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4757
52 4699 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4758
53 4700 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4759
54 4701 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4760
55 4702 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4761
56 4703 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4762
57 4704 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4763
58 4705 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4764
59 4706 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4765
60 4707 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4766
61 4708 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4767
62 4709 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4768
63 4710 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4769
64 4711 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4770
65 4712 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4771
66 4713 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4772
67 4714 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4773
68 4715 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4774
69 4716 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4775
70 4717 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4776
71 4718 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4777
72 4719 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4778
73 4720 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4779
74 4721 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4780
75 4722 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4781
76 4723 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4782
77 4724 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4783
78 4725 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4784
79 4726 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4785
80 4727 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4786
81 4728 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4787
82 4729 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4788
83 4730 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4789
84 4731 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4790
85 4732 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4791
86 4733 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4792
87 4734 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4793
88 4735 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4794
89 4736 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4795
90 4737 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4796
91 4738 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4797
92 4739 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4798
93 4740 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4799
94 4741 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4800
95 4742 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4801
96 4743 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4802
97 4744 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4803
98 4745 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4804
99 4746 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4805
100 4747 gttatataaactacatacaaaagatctacatctagcccacactctctacaaatgataaa 4806

```

```

RESULT 10
10 254874 standard: DNA: 6555 BP.
A0 254874

```

```

DT 27-JUN-1994 (first entry)
DE Plasmid pVE144 used in the construction of plant maintainer gene.
KW Maintainer gene: sterile; sterility: homogeneous population; hybrid;
   seed; fertility restorer gene; pollen lethality gene; ds.
OS Synthetic.
PB Key.
FT Location/Qualifiers
   FT misc_feature 1..396
      FT /label= pUC18.
      FT /note= "gene derived 3042 bp."
      FT /tag= b
      FT /tag= b
      FT /label= 3' nos.
      FT /note= "3' regulatory sequence containing the
      polyadenylation site derived from
      Arabidopsis T-DNA napaline synthase gene."
      FT misc_feature 752..1024
      FT /tag= c
      FT /label= Barstar.
      FT /note= "Coding region of the barstar gene of
      Bacillus amyloliquefaciens."
      FT promoter 1025..1607
      FT /tag= d
      FT /label= TA29.
      FT /note= "Promoter derived from TA29 gene of Nicotiana
      tabacum."
      FT promoter 1608..2440
      FT /tag= e
      FT /note= "4684 promoter sequence derived from
      cauliflower mosaic virus isolate Cabbage 11."
      FT cds 2441..3256
      FT /tag= f
      FT /note= "Coding region of the tobacco
      phosphotransferase gene of TMV."
      FT misc_signal 3257..4315
      FT /tag= g
      FT /note= "3' regulatory sequence containing the
      polyadenylation site derived from
      Arabidopsis T-DNA octopine synthase gene."
      FT misc_feature 4316..6555
      FT /tag= h
      FT /label= pUC18.
      FT /note= "pUC18 derived sequence."
      WC9325695-A.
      23-DEC-1993.
      PD 11-JUN-1993; E01489.
      PR 12-JUN-1992; US-849072.
      PR 03-NOV-1992; US-970840.
      PA (PLB2 ) PLANT GENETIC SYSTEMS NV.
      PI Leemans J., Williams M.
      DR WPI: 94-007552/01.
      PT Maintainer gene for maintenance of male-sterile plants
      PT comprises fertility-restorer gene and pollen-lethality gene.
      PT Example 2: Page 50-54; 87pp; English.
      TS A maintainer gene of plants, pref. a foreign chimeric gene,
      CC comprises (a) a fertility restorer gene which comprises a fertility
      CC restorer DNA and (ii) a restorer promoter capable of
      CC directing the expression of the fertility restorer DNA and (b) a
      CC pollen lethality gene that is selectively expressed in microspores
      CC and/or pollen of the plant to prevent the production of functional
      CC pollen and which comprises (i) a pollen lethality DNA and (ii) a
      CC pollen specific promoter capable of directing expression of the
      CC pollen lethality DNA. Plants transformed with this DNA (maintainer
      CC plants) can be used to maintain a homogeneous population of male-
      CC sterile plants for the production of hybrid seed. This plasmid
      CC contains DNA encoding Barstar (the fertility restorer DNA), the TA29
      CC tapetum-specific promoter (the restorer promoter), nos as a marker
      CC DNA and the 35S3 promoter to express the marker DNA. It was used
      CC alongside another plasmid (pVE108, described in file 954875) in the
      CC construction of a plant transformation vector comprising a
      CC maintainer gene as described.
      CC Sequence 6555 BP; 1690 A; 1612 C; 1583 G; 1670 T;
      SQ

```

Query Match 40.7% Score 2241 DB 9 Length 6555.
Best Local Similarity 100.0% P-Id No 6 00e+00.
Matches 2241 Conservative 0 Mismatches 0 Indels 0 Gaps 0

Db	4315	caagcttgcgttaatactgctcacaactgttccctgtgtgaataattgttatccgctcaca	4374
QY	3271	CAAGCTTGGCGTAATCATGGTCATAGCTGTTCCCTGTGTGAAATGTTATCCGCTCACA	3330
Db	4375	ttccacacacatacagccggaagcacaagaatgtataaaccttgggtgccttaataagta	4434
QY	3331	TTCACACACATACAGCGCGAAGATATAAGAGTAAAGGCTTGGGTCCTTAATGAGTGA	3390
Db	4435	gtaactcacataattgctgttgcctcactgcctgccttccagtcgggaacactgtcgt	4494
QY	3391	GTAATATACATTAATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG	3450
Db	4495	scgaactcacataataatcgcccaacgcgcggggaagagcgggttgggtattgggcct	4554
QY	3451	GGCAGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	3510
Db	4555	cttccgcttccctgcctcactgactgcctgcctgcctgcctgcctgcctgcctgcct	4614
QY	3511	CTTCCGTTCCCTGCCTACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3570
Db	4615	caactcaactcaaggcggtaactgattatccacacaaatcagggtataacgcaggaaaga	4674
QY	3571	CAGCTCACACAAAGGCGGTAATAGGTAATAGGTAATAGGTAATAGGTAATAGGTAAT	3630
Db	4675	acatgtgaacaaagccagcaaaagccaggaacccgtataaagccgcgttgcgtgcgt	4734
QY	3631	ACATGTGACCAAAAGGCGGTAATAGGTAATAGGTAATAGGTAATAGGTAATAGGTAAT	3690
Db	4735	ttttccataggctccccccttcgcagagcatcacaaaaatcgagcgtcaactcagaggt	4794
QY	3691	TTTTCCATAGGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3750
Db	4795	gacgaacacagacacataaagatacagggctttcccccctggaagctccctgcgtc	4854
QY	3751	GACGAACACAGACACATAAAGATACAGGTAATAGGTAATAGGTAATAGGTAATAGGTAAT	3810
Db	4855	gcctcctcttccgacccctgcgcgttaacggatacctgtccgccttctcccttcggaaa	4914
QY	3811	GCTCTCCTTCCGACCCCTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3870
Db	4915	gctgacgcttttccatgctcagcgttagtatctcagttcagttcagttcagttcagtt	4974
QY	3871	GCTGACGCTTTTCCATAGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3930
Db	4975	ccagctgagctgtgtgagcaaccccgcttcagcccaacgcgtgcgcttatccgcta	5034
QY	3931	CCAGCTGAGCTGTGTGAGCAACCCCGCTTCAGCCCAACGCCTGCTGCTGCTGCTGCTG	3990
Db	5035	actatgcttttgcgtccaaacccggttaagcacagactatcgccactggcagcagcactg	5094
QY	3991	ACTATGCTTTTGCCTCAACCCCGCTTAAGCACAGACTATCGCCACTGGCAGCAGCAGT	4050
Db	5095	gtaacangattagcagcagaggtatgtagggggtgctcacaagttcttgaagtggtgagc	5154
QY	4051	GTAACANGATTAGCAGCAGAGGTATGTAGGGGTGCTCACAAGTTCTTGAAGTGTGTC	4110
Db	5155	ctaactcacgctacataaggaagcagttatgtgtatcgtcgtcgtcgtcgtcgtcgtc	5214
QY	4111	CTAACTCACGCTACATAAGGAAGCAGTTATGTGTATCAGTGTGCTGCTGCTGCTGCTG	4170
Db	5215	cttctggaacaaaggt	5274
QY	4171	CTTCTGGAACAAAGGT	4230
Db	5275	gtttttttgttttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	5334
QY	4231	GTTTTTTGTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	4290
Db	5335	tgaatcttctcaggggtgtgagcgtcagtggaacaaactcaactcaggttaagggattgg	5394

QY	4291	IGATCTTTTCTAGGGGTCTGAGGCTCAGTGGAGCAAGAAATACAGTAAAGGATTTTGG	4350
Db	5395	tcagagattatcaaaaagagatttccacctagatctcttttaaatataaatgaatttta	5454
QY	4351	TCATGAGATTATCAAAAAGGATTTTCCCTAGATCTCTTTTAAATTAATAATGAAGTTTA	4410
Db	5455	aatacaactaaagtatatagataaacttgggtctcaacttaccacttcaacttcaact	5514
QY	4411	AATCAATCTTAAGTATATATAGTAACCTTGGTCTGACAGTATACCAATGCTTAAATAGTG	4470
Db	5515	aggcacctatctcagcagctgtctatttcttccatctccttagttgctgactcccggtcg	5574
QY	4471	AGGCACCTATCTCAGCGATCTGCTATTCTGCTCATCTAGTGGTSACTCCGAGTGG	4530
Db	5575	tgtagataactacagatcagcggagggttaccatctggcccccaatctgcaatataccgc	5634
QY	4531	TGTAGATAACTAGATACGGGAGGGGTATATCTGCGGCGCACTGGTCAATGATACGG	4590
Db	5635	gagaccacacgtcaccggctcccaatttaacacaaataaacacacagcagcagcagcag	5694
QY	4591	GAGACCCACGCTCACCGGCTCCAGATTATACGAAATAAACTAAGCAAGCAAGCAAGCA	4650
Db	5695	agcgagaagtgtctcgtcaacttttccgctccctccactcttattatttggttgcggg	5754
QY	4651	AGCGCAGAAAGTGTCTGCAACTTTATCCGCTCCATCTCAAGTATTAATTTGTTGGCGG	4710
Db	5755	aaqctagaagtaagtagtcccgagtttaagtgtggcgaacggtgtgtgcatgctacag	5814
QY	4711	AAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTGGCAAGGTTGTTGCAATTTGCTACAG	4770
Db	5815	gcactgtgtgtcagcgtcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	5874
QY	4771	GCATCGTGGTGCAGCGTCTGCTTTGCTATGCTTTCATTCAGCTACGGTTCGCAACGAT	4830
Db	5875	caaggcaggtatcatgctcccccatgttgtgcaaaaaagcggttagctccttccctc	5934
QY	4831	CAAGCGGAGTTATATGATCTCCCATGTTGTTGTCAGAAAAAGGTTAGTTCCTTTGCTGCT	4890
Db	5935	cgactgtgtcgaagaagtagtggcgcaggtgttatcactcaggttatgagcagcagcag	5994
QY	4891	CGATCGTGTGCAGAAAGTAAAGTGGCGGCAATGTTATCTCAATATGCTTATGCTATGCT	4950
Db	5995	ataattcttcttctcagcctccatcagatcctttctgtgactgactgactgactgact	6054
QY	4951	ATAATTCTCTTACTGTCATGCCATCTGCTAAAGTCTTTTCTGTTGCTGCTGCTGCTGCT	5010
Db	6055	ccaagctcattctgaagaatagttatggcgagcagcagcagcagcagcagcagcagcag	6114
QY	5011	CCAAGCTCATTTGAGAAATAGTATAGCGGCGAGCGAGTTGCTTTGGCGGCGCTCAATAG	5070
Db	6115	gggataataaccgcccacatagcagaactttaaaagtgtcctcactggtgaaagcgttctt	6174
QY	5071	GGGATAATACCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAACGTTCTT	5130
Db	6175	ggggcgaaaaactctcagatcttaccgctgttgaactccacttccatgactgactgact	6234
QY	5131	CGGGCGAAAAACTCTCAAGGATCTTACGCTGTTGCTATCTCAATCTCAATCTCAATCT	5190
Db	6235	gtgaccccaactgacttccagcatttttacttttccacacgcttttccggttgaagcaaaa	6294
QY	5191	GTGACCCCAACTGATCTTCAGCATCTTTTAACTTTTCACTAGCTTTTCTGGGCTGAGCAAAA	5250
Db	6295	caggaaagcaaaaagcgcgaacaaaggaataagggcgacacaggaatgttgaactcga	6354
QY	5251	CAGAAAGCAAAAATTCGCGCAAAAAAGGAAATAAGGCGACACGCAAAAGTGTGAATCTCA	5310
Db	6355	tactctctcttttcaatattattgaagcatttatcaggttatttctctcatgagcgat	6414
QY	5311	TACTCTCTCTTTTCAATATTATTGAAGCAATTTATCAAGGTTTATTTGCTCATGAGCGAT	5370
Db	6415	acatatttgaatgatttagaaaaataaataaataaggggttccgcgcagcaatctcccgaa	6474

PI Martin SL, Weiss AS;
DP WP: 94-263633/2
PT Synthetic polynucleotide(s) - encode recombinant tropoelastins
PI and variants
PS Disclosure: Page 48; 77pp; English.
CC This non-fusion vector has the human synthetic tropoelastin gene
CC inserted into an NcoI site and uses phage lambda promoter-driven
CC expression to permit initiation from the start codon of the
CC synthetic gene. The expressed product is susceptible to
CC hydrolytic breakdown of the crosslinks. Such material may be
CC useful in e.g. surgical applications, where the gradual loss of
CC material over time is intended.
SQ Sequence 4045 BP; 1045 A; 994 C; 956 G; 1049 T;
Query Match 40.6%; Score 2239; DB 12; Length 4045;
Best Local Similarity 100.0%; Pred. No. 0.0e+00;
Matches 2240; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 399 gacaaaggccctcgatagcctatttttataggtaatgcataataatggttt 458
Cp 5511 GACCAAGGCGCTCGTATAGCGCTATTTTATAGGTTAATGTCATGATAAATGGTT 5452
Db 459 cttagacgtcaggtggcacatttcggggaaatattgcgcgaacccctatttggtttttt 518
Cp 5451 CTTAGACGTCAAGTGGTACATTTTGGGGAATGTCGCGGAACCCCTATTGTGTTATTT 5392
Db 519 tctaaatcacattcaaatatgtatccgctcatgagacataaacctctgataaatgcttcaat 578
Cp 5331 TCIAAATACATTCAAATATGTATCGGTGTCATGAGACAAATAACCTGTATAATCGTTCAAT 5332
Db 579 aatattcaaaagaagaagatgagtattcaacatttcctgctgcgccttattccctttt 638
Cp 5331 AATATTGAAAAGGAGAGATGATGATGATTCACCAATTTCCCGTTCGCCCTTATTCCCTTT 5272
Db 639 ttgcgcatcttgctccctgtttttactcaccagaaacgcgtggtaagaataaagatg 698
Cp 5271 TTGAGCAATTTTGGCTTCCCTTTTTCCTACCTAGAAACGCTGGTGAAAGTAAAGATG 5212
Db 699 ctgaagatcagttgggtgcacgagtggtttacatcgcaactggatctcaacagcggttaaga 758
Cp 5211 CTGAAGATCACTTGGTGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAG 5152
Db 759 tccgtgagagtttcgcccgaagaacagttttccaatgataagcaacttttaaggttctgc 818
Cp 5151 TCCTTTGACAGTTTTCGCTCCGCGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGC 5092
Db 819 tatctggcgcggtattatcccgatttagcgcggggaagagcaactcggtcgccgcatac 878
Cp 5091 TATGTGGCGGGTATTATCCGGTATTGCGCGGGGGAAGAGCACTCGTCCGCCGATAC 5032
Db 879 actattctcagaatgacttggttgagttactcaccagttcacagaaaagcatcttcacggatg 938
Cp 5031 ACTATTCTCAATGACTTGGTTGAGTATCACCAGTCAACAGAAAGCATCTTACGGATG 4972
Db 939 gcatgaagtaagaagaattatgcaatgctgccataaacatgagtgataaacactgcgggcca 998
Cp 4971 GATGACAGTAAGAAGAT 4912
Db 999 acttacttctcaacacatcggagagccggaagagatgaacccgttttttgcacaaactgg 1058
Cp 4911 ACTTACTTCTGACAAAGATGAGAGACGCGAAGAGCTAACCGCTTTTTCGACAAACATGG 4852
Db 1059 ggaatcatgataactcgccttgatcgttgggaacccggagctgaatgaagccataccaaag 1118
Cp 4851 GGGATCATGTAATCTGCTGCTGATGCTTGGGAACCGGAGCTGAATGAAGCCATACCAAAAG 4792
Db 1119 acgaactgacacacatccttatacaatgcaacaaacagtttcgcgaacacttaactg 1178
Cp 4791 ACAGGCGTGAACACACAGATGCTGAGCAATGCGAAACAGCTTCCGGAACACTATTAAGTC 4732
Db 1179 qcgaactacttactctagcttcccggaacaaataatagactgagtgagcgagataaag 1238
Cp 4731 CGGAACACTTACTTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAG 4672

Db 1239 ttacagggacacacttctgaggggctgg 1238
Cp 4671 TTGCAGGAGCACCTTCGCGCTCGGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGG 4612
Db 1299 gagccggtgagcgtgggtctcgcgcgtatcattgcagcactggggcccaaatgataagcctt 1358
Cp 4611 GAGCTCGTGAAGGCTGGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGG 4552
Db 1359 cccgtatcgtgattatctacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1418
Cp 4551 CCCGTATCGTATATCTTACACAGCAGGAGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4492
Db 1419 agatcgctgagatagtgccctcaactgatttaagcattgtaactatcaacacaaatttact 1478
Cp 4491 AGATCGCTGAGATAGTCCCTTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4432
Db 1479 catatactttagattgatttaaaacttcttttaaaaggagatctagatgaaga 1538
Cp 4431 CATATATCTTTAGATTGATTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGA 4372
Db 1539 tcccttttgataatctcatgacacaaatcccttaacgtaggttttcgttcacgtgaggtt 1598
Cp 4371 TCCCTTTTGATAATCTCATGACCAAAATCCCTTAAAGTCAAGTTCGCTTCAGTCAAGTTCAGGCT 4312
Db 1599 cagaccccgtagaagaatcacaagatctcttgaagatcccttttctgcgcgttaactt 1658
Cp 4311 CASAGCCGCTAGAAAAGATCAAGGATCTCTTTCAGATCTCTTTTTCGTGCGCTTAATCT 4252
Db 1659 gctgtttgcaacaaaaaacacccgctaccagcggtgttttttgcgcgactcaagagc 1718
Cp 4251 GCTGCTTGCACAAACAAAAAACACCGCTACGAGGCTGGTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 4192
Db 1719 taccacactcttttccgaaggttaactgcttcacgacagcagcagcagcagcagcagcagcagcagc 1778
Cp 4191 TACCACACTCTTTTTCGAGGTTAATCTGCTTCAGCAGAGGCTGCTTCAGCAGAGGCTGCTTCAGCAG 4132
Db 1779 ttctagtgtagcgttagttaggccaacccttcaagaactctgtacacagcagcagcagcagcagc 1838
Cp 4131 TTCTAGTGTAGCGGTAGTTAGGCGACCACTTCAGAACTCTGTAGCAGCAGCAGCAGCAGCAGCAG 4072
Db 1839 tccctctactaatctctgttaccagtggtgctgctgacagtgagcagcagcagcagcagcagcagcagc 1898
Cp 4071 TCGCTCTGCTAATCTCTTACAGTGGTTCGCTTCAGCAGAGGCTGCTTCAGCAGAGGCTGCTTCAGCAG 4012
Db 1899 gcttgagctcaagacgataattacagataagagcagcagcagcagcagcagcagcagcagcagcagc 1958
Cp 4011 GGTGGACTCAAGACGATAGTTTACCGGATAGAGGCTGCTTCAGCAGAGGCTGCTTCAGCAGAGGCT 3952
Db 1959 cgtgcacacagcccgcttgagcgaacgacctacacccgaactgagatcactacagcgtg 2018
Cp 3951 CTTGACACAGGCTCAGTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3892
Db 2019 agcatgtgaaagcgcacgcttcccgaaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 2078
Cp 3891 AGCATTGAAAGCGCTACGCTTCGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3832
Db 2079 gcaggttcggaacagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2138
Cp 3831 GCAGGCTCGGAACAGGAGCGCACAGGAGCTTCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3772
Db 2139 atagtcctgtcggttttcgccaacctctgaactgagcgtcagatcttttctgagatgagcgtgag 2198
Cp 3771 ATAGTCTGTGCGGTTTCGCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 3712
Db 2199 gggggcgagcctatggaaaaaacgcac 2258
Cp 3711 GGGGCGGAGCGCTATGAAAAAACGACCAAGTGGGCTTTTAAAGGCTTCGCTTCAGTTCAGTTCAGTTCAG 3652
Db 2259 gctggccttttgcctcactgttcttctgaggtatccctgagattctatgataaacta 2318
Cp 3651 GCTGGCTTTTCTCAGATGTTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 3592

Db 2086 atgaattatcaaaagagatctccactagatccttttaataataaataaagattttaa 2145
 QY 4353 ATGACATATCAAAAAGGATCTTCACCTAGATGCTTTTAAATTAATAATGAAGTTTAAA 4412
 Db 2146 tcaatcaaatatataatgaataaacttggtctgacaggtaccatgcttaatacagtgag 2205
 QY 4413 TCAATCAAAATATATATGATTAAGTAAAGTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG 4472
 Db 2206 gacactatctcagcagatctctctattctcattctccatccatagttgctgactcccctgctg 2265
 QY 4473 GCACCTATCTAGCCATCTGTCTCTATTCTGTTCTATCTCATCTAGTTGCTGACTCCCTGCTG 4532
 Db 2266 tagataactcagatcagcagagaggtaccatctctggcccatgctgcaatgataccgcga 2325
 QY 4533 TAGATAACTATGATACGAGAGAGGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4592
 Db 2326 gaccacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2385
 QY 4593 GACCCACGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4652
 Db 2386 cgcacaaatggtcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2445
 QY 4653 GCGCAAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4712
 Db 2446 gctagatgaagttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2505
 QY 4713 GCTAGACTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4772
 Db 2506 atcgtatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2565
 QY 4773 ATCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4832
 Db 2566 agcagatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2625
 QY 4833 AGCGAGTTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4892
 Db 2626 atcgtatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2685
 QY 4893 ATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4952
 Db 2686 aattcttactgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2745
 QY 4953 AATTCTCTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5012
 Db 2746 aattcttactgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2805
 QY 5013 AAGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5072
 Db 2806 gataatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2865
 QY 5073 GATAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 5132
 Db 2866 gacgcaaaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2925
 QY 5133 GGGCGAAAACCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 5192
 Db 2926 gacgcaaaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2985
 QY 5193 GCACCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5252
 Db 2986 gaaatgcaaaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3045
 QY 5253 GAAGGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5312
 Db 3046 ct 3105
 QY 5313 CT 5372
 Db 3106 atattggaatgatttagaataataaataataggggttcgcgcgcacatttccccgcgaaga 3165
 QY 5373 ATATTGGAATGATTAGAAAAATAAACAATAAGGGGTTCGCCGCGACATTTCGCCCGGAAGA 5432

Db 3166 gtacacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3225
 QY 5433 GTGCCACTCTGAGCTGTATGAAGAGCTATTAATCATGACATTAACCTAIAAAAAAIAAGGCT 5492
 Db 3226 atcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3244
 QY 5493 ATCAGCAGGCGCCTTTCGTC 5511
 RESULT 15
 ID 169188 standard: DNA: 4118 BP.
 AC 169188;
 DT 26-FEB-1998 (first entry)
 DE Construct pGEM-HTR containing RNA component of human telomerase.
 KW Human telomerase; quantification; tumour cell; pGEM-HTR;
 KW detection; micrometastasis; diagnosis; lymphoblastoma; leukaemia;
 KW teratocarcinoma; melanoma; carcinoma; cancer; tumour; neuroblastoma;
 KW rhabdomyosarcoma; leiomyosarcoma; lymphoma; RNA component; ss.
 OS Homo sapiens.
 CS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 12..975 FT
 FT /*tag= a
 FT /note= *cDNA from RNA component of human telomerase*
 FT WO9718322-A2.
 PN 22-MAY-1997.
 PD 14-NOV-1996; D02183.
 PF 16-NOV-1995; DE-042795.
 PR (DAHM/) DAHM M.W.
 PA Dahm MW;
 PI WPI: 97-289298/26.
 DR Quantifying tumour cells in body fluid - by measuring RNA component
 FT of telomerase after amplification, especially useful for early
 PT diagnosis of metastasis
 PS Example 5; Fig 5; 46pp; German.
 CC The present sequence is the construct pGEM-HTR, which comprises the
 CC transcription vector pGEM-13zf(+) and the cDNA from the RNA
 CC component of human telomerase. The construct was used in the
 CC development of a novel method for quantifying tumour cells in a
 CC body fluid. The method comprises specific amplification of the RNA
 CC component of telomerase, and measuring the amount of amplified
 CC nucleic acid. At least 1, preferably all 3 (769173-75) standard
 CC nucleic acids are present (at different concentrations) and are
 CC co-amplified with telomerase DNA. The amplification products are
 CC detected directly or via a label or by hybridisation with a
 CC labelled oligonucleotide (769184-87, 1 for each standard and 1 for
 CC telomerase), and the amount of telomerase product compared with
 CC that from the standards.
 CC The method can be used to detect tumour cells, specifically
 CC micrometastases, in a body fluid, particularly for the early
 CC diagnosis of metastatic spread and for monitoring tumour therapy.
 CC Typical tumour cells that can be quantified are micrometastases,
 CC T cell lymphoblastoma or leukaemia (chronic myelogenous, and
 CC chronic or acute lymphatic leukaemia), teratocarcinoma, melanoma,
 CC carcinoma of lung, liver or prostate, cancer of colon or breast,
 CC kidney, brain or adrenal tumours, neuroblastoma, rhabdomyosarcoma,
 CC leiomyosarcoma and/or lymphoma.
 CC Sequence 4118 BP, 531 A, 1032 C, 1094 G, 1001 T;
 SQ
 Query Match 40.88; Score 2235, DB 35; Length 4118-
 Best Local Similarity 99.94; Pred. No. 0.00e+00;
 Matches 2237; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 1006 agcttgccgctaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1065
 QY 3273 AGCTTGGCGTAATCATGCTGCTATAGCTGTTCCTGTGTGAAATTTTATGCGCTCAAAAT 3332
 Db 1065 ccaacacacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1125
 QY 3333 CCACACAAACATACGAGCGCGAAGCATAAAGTGTAAAGCTTGGGCTTCTCTAATAGTGAAGC 3392
 Db 1126 taactacatgaattcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1185

W P S R E H

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_on a - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Mar 15 04:49:36 1999. MasPar time 7003.29 Seconds
Tabular output not generated
1409.192 Million cell updates/sec
Title: >US-09-020-716-4
Description: (1,551) ftm US09020716.seq
Perfect Score: 5511
N.A. Sequence: 1 TCGCGCTTTCGTGATGAC..... TATCAGAGGCGCTTTCGTC 5511
Comp: AGCGCGCAAGCAGCTACTG ATAGTGTCTCGGCAAGGACG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109

5.gb-est1 6.gb-est10 7.gb-est11 8.gb-est12 9.gb-est13
10.gb-est14 11.gb-est15 12.gb-est16 13.gb-est17
14.gb-est18 15.gb-est19 16.gb-est20 17.gb-est21
18.gb-est22 19.gb-est23 20.gb-est24 21.gb-est25 22.gb-est26
23.gb-est27 24.gb-est28 25.gb-est29 26.gb-est30 27.gb-est31
28.gb-est32 29.gb-est33

Statistics: Mean 12.756; Variance 2.337; scale 5.464

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	564	12.0	711	26	AG004608 Homo sapiens genomic D	0.00e+00
2	564	12.0	711	27	AG010947 Homo sapiens genomic D	0.00e+00
3	557	11.9	721	27	AG011001 Homo sapiens genomic D	0.00e+00
4	557	11.9	721	25	AG004552 Homo sapiens genomic D	0.00e+00
5	530	11.4	698	26	AG003787 Homo sapiens genomic D	0.00e+00
6	530	11.4	698	27	AG009976 Homo sapiens genomic D	0.00e+00
7	531	11.4	718	26	AG004363 Homo sapiens genomic D	0.00e+00
8	531	11.4	718	27	AG010489 Homo sapiens genomic D	0.00e+00
9	524	11.3	696	27	AG009765 Homo sapiens genomic D	0.00e+00
10	524	11.3	696	26	AG003576 Homo sapiens genomic D	0.00e+00
11	520	11.2	720	25	AG006052 Homo sapiens genomic D	0.00e+00
12	520	11.2	720	28	AG013858 Homo sapiens genomic D	0.00e+00
13	519	11.2	747	25	AG007052 Homo sapiens genomic D	0.00e+00

C	14	604	11.0	692	27	AG040352	CIT-HSP-237K01.TF CIT	0.00e+00
	15	603	10.9	695	26	AG002183	Homo sapiens genomic D	0.00e+00
	16	594	10.8	640	27	AG004298	CIT-HSP-2382J1.TF CIT	0.00e+00
	17	593	10.8	703	26	AG001761	Homo sapiens genomic D	0.00e+00
	18	594	10.8	748	17	AI124381	1.53 Drosophila 8-12 h	0.00e+00
	19	589	10.7	452	28	AG008233	CIT-HSP-237M3.TF CIT	0.00e+00
	20	590	10.7	690	27	AG009464	Homo sapiens genomic D	0.00e+00
	21	584	10.6	466	27	AG079096	CIT-HSP-235E54.TF CIT	0.00e+00
	22	579	10.5	594	26	C17B6	Ciona intestinalis gen	0.00e+00
	23	577	10.5	682	28	AG014394	Homo sapiens genomic D	0.00e+00
	24	568	10.3	642	28	AG112213	CIT-HSP-237L22.TF CIT	0.00e+00
	25	558	10.1	615	28	AG112424	CIT-HSP-237F2.TF CIT	0.00e+00
	26	551	10.0	604	26	AG002706	Homo sapiens genomic D	0.00e+00
	27	551	10.0	634	27	AG038010	CIT-HSP-238F2.TF CIT	0.00e+00
	28	550	10.0	653	27	AG077534	CIT-HSP-235K1.TF CIT	0.00e+00
	29	545	9.9	633	27	AG040787	CIT-HSP-233K2.TF CIT	0.00e+00
	30	546	9.9	651	27	AG057693	CIT-HSP-234M2.TF CIT	0.00e+00
	31	541	9.8	637	27	AQ077300	CIT-HSP-2365E1.TF CIT	0.00e+00
	32	537	9.7	626	28	AQ111342	CIT-HSP-237M2.TF CIT	0.00e+00
	33	535	9.7	682	27	AQ074693	CIT-HSP-2301L23.TF CIT	0.00e+00
	34	526	9.5	540	27	AG074307	CIT-HSP-2382N1.TF CIT	0.00e+00
	35	526	9.5	594	26	AG067826	Homo sapiens genomic D	0.00e+00
	36	526	9.5	600	28	AQ111411	CIT-HSP-237F2.TF CIT	0.00e+00
	37	523	9.5	629	27	AG011367	Homo sapiens genomic D	0.00e+00
	38	526	9.5	743	26	AG002826	Homo sapiens genomic D	0.00e+00
	39	517	9.4	623	27	AQ081172	CIT-HSP-2357C1.TF CIT	0.00e+00
C	40	518	9.4	647	27	AG058943	Homo sapiens genomic D	0.00e+00
	41	514	9.3	612	27	AG080596	CIT-HSP-2382D1.TF CIT	0.00e+00
	42	510	9.3	644	27	AG012911	Homo sapiens genomic D	0.00e+00
	43	510	9.3	664	26	AG005286	Homo sapiens genomic D	0.00e+00
C	44	509	9.2	705	27	AG009803	Homo sapiens genomic D	0.00e+00
C	45	509	9.2	705	26	AG003614	Homo sapiens genomic D	0.00e+00

ALIGNMENTS

RESULT 1	AG004608	711 bp	DNA	GSS	29-JAN-1998
LOCUS	Homo sapiens genomic DNA, 21q region, clone: S594BG38, genomic survey sequence.				
DEFINITION	AG004608				
ACCESSION	AG004608				
NID	Q2822058				
KEYWORDS	GSS.				
SOURCE	Homo sapiens DNA, clone: S594BG38.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 711)				
AUTHORS	Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
TITLE	Eukaryotic, Metazoa, Chordata, Vertebrata: Mammalia: Eutheria: Primates, Catarrhini, Hominoidea: Homo.				
JOURNAL	Published Only in DataBase (1998) In press				
REFERENCE	2 (bases 1 to 711)				
AUTHORS	Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.				
	Masahira Hattori, Kitasato University, Department of Science, 357 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 223, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp. Tel: 0427-78-9732. Fax: 0427-78-9561)				

FEATURES

source	Location/Qualifiers
1..711	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="21"	
/clone="S594BG38"	
/map="21q"	
BASE COUNT	169 a 188 c 172 g 192 t
ORIGIN	
Query Match	12.0%; Score 664; DB 26; Length 711;
Best Local Similarity	99.6%; Pred. No. 0.00e+00;
Matches	682; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

```

Db 28 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 87
QY 4533 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 4591

Db 88 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147
QY 4592 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4651

Db 148 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
QY 4652 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4711

Db 208 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
QY 4712 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4771

Db 268 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
QY 4772 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4831

Db 328 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
QY 4832 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4891

Db 388 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 4892 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4951

Db 448 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
QY 4952 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5011

Db 508 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
QY 5012 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5071

Db 568 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 5072 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5131

Db 628 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 686
QY 5132 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5190

Db 687 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 5191 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5215

RESULT 2 AG010947 711 bp DNA GSS 28-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone S594HG38, genomic
DEFINITION survey sequence.
ACCESSION AG010947
NID g334775
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 711)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 711)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1998) to the EMBL/GenBank/GenBank databases
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@shimizu.ims.kitakyu.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
LOCATION/Qualifiers

Db 28 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 87
QY 4533 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 4591

Db 88 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147
QY 4592 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4651

Db 148 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
QY 4652 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4711

Db 208 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
QY 4712 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4771

Db 268 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
QY 4772 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4831

Db 328 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
QY 4832 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4891

Db 388 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 4892 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4951

Db 448 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
QY 4952 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5011

Db 508 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
QY 5012 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5071

Db 568 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 5072 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5131

Db 628 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 686
QY 5132 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5190

Db 687 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 5191 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5215

RESULT 3 AG011001 721 bp DNA GSS 29-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone S796LX91, genomic
DEFINITION survey sequence.
ACCESSION AG011001
NID g3357935
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1998) to the EMBL/GenBank/GenBank databases
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@shimizu.ims.kitakyu.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
LOCATION/Qualifiers

```

```

source 1. 711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="S594HG38"
/map="21q"

BASE COUNT 169 a 188 c 172 g 182 t
ORIGIN

Query Match 12.0% Score 664; DB 27; Length 711;
Best Local Similarity 66; pred No 0.00e+00;
Matches 682; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

Db 28 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 87
QY 4533 TAGATAAATACGATACAGAGAGAGATTAATGATGAGGATGAGTGAATGATGAGGAGG 4591

Db 88 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 147
QY 4592 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4651

Db 148 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
QY 4652 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4711

Db 208 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
QY 4712 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4771

Db 268 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
QY 4772 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4831

Db 328 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
QY 4832 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4891

Db 388 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
QY 4892 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4951

Db 448 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
QY 4952 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5011

Db 508 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
QY 5012 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5071

Db 568 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
QY 5072 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5131

Db 628 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 686
QY 5132 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5190

Db 687 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 5191 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5215

RESULT 3 AG011001 721 bp DNA GSS 29-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clone S796LX91, genomic
DEFINITION survey sequence.
ACCESSION AG011001
NID g3357935
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 721)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 721)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1998) to the EMBL/GenBank/GenBank databases
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@shimizu.ims.kitakyu.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
LOCATION/Qualifiers

```


REFERENCE		1 (bases 1 to 721)	
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
TITLE	Homo sapiens genomic DNA, chromosome 21q		
JOURNAL	Published Only in Database (1998) In press		
REFERENCE		2 (bases 1 to 721)	
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.		
	Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 229, Japan		
	(E-mail:hattori@erc.ims.u-tokyo.ac.jp; Tel:0427-78-9732, Fax:0427-78-9561)		
FEATURES	Location/Qualifiers		
source	1..721		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="21"		
	/clone="879G12X91"		
	/map="21q"		
BASE COUNT	195 a 163 c 175 g 179 t	9 others	
ORIGIN			
Query Match	11.98;	Score 657;	DB 27; Length 721;
Best Local Similarity	98.74;	Pred. No. 0.00e+00;	
Matches	683; Conservative	0; Mismatches	5; Indels 4; Gaps 4;
Dbb	32 CATCTACGGATGGTATGATGAAGAAATATGTCAGTGCTGCCATAACCATTGAGTGAT	91	
Cp	4984 CATCTACGGATGGTATGATGAAGAAATATGTCAGTGCTGCCATAACCATTGAGTGAT	4925	
Dbb	92 AACACTCGGGCCCAACTTACTTCTGACAAACGATGAGAGAGGTAAACCGGTTTT	151	
Cp	4924 AACACTCGGGCCCAACTTACTTCTGACAAACGATGAGAGAGGTAAACCGGTTTT	4865	
Dbb	152 TTGCACAACATGGGGATCATGTAACCTGCGCTTGATCGTTGGAAACCGGAGCTGAATGAA	211	
Cp	4864 TTGCACAACATGGGGATCATGTAACCTGCGCTTGATCGTTGGAAACCGGAGCTGAATGAA	4805	
Dbb	212 GCATATCAATGACACAGGCGTAGAATCAATGAGTGGTGTAGTAATGCAACAACTGGCG	371	
Cp	4804 GCATATCAATGACAGGCGTAGAATCAATGAGTGGTGTAGTAATGCAACAACTGGCG	4745	
Dbb	272 AAACATTAACTGGCAACTTACTTACTTASCTTCGGCGCAACAATTAATAGACTGGATG	331	
Cp	4744 AAACATTAACTGGCAACTTACTTACTTASCTTCGGCGCAACAATTAATAGACTGGATG	4685	
Dbb	332 GAGCGGATAAATTTGACGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCTTTAT	391	
Cp	4684 GAGCGGATAAATTTGACGACCACTTCTGCGCTCGGCGCTTCGCGCTGGCTTTAT	4625	
Dbb	392 GCTGATAAATCTGAGCGGTGAGCGTGGTCTCTCGGGTATCATTCAGCACCTGGGGCCA	451	
Cp	4624 GCTGATAAATCTGAGCGGTGAGCGTGGTCTCTCGGGTATCATTCAGCACCTGGGGCCA	4565	
Dbb	452 GATGGTAAGGCTCTCTGATATCTGATGATATCTACACACGCGGGAGTCAGCAACTTGGAT	511	
Cp	4564 GATGGTAAGGCTCTCTGATATCTGATGATATCTACACACGCGGGAGTCAGCAACTTGGAT	4505	
Dbb	512 GAACGAAATACACAGATCGGTACAGTAGTGTCCTCCTACTGATTAAGC-TTGGTAACCTGCA	570	
Cp	4504 GAACGAAATACACAGATCGGTACAGTAGTGTCCTCCTACTGATTAAGC-TTGGTAACCTGCA	4445	
Dbb	571 GACCAAGTTTACTC-TATATACTTTAGATTGATTTAAAACCTCATTTTTTAATTTAAAGG	629	
Cp	4444 GACCAAGTTTACTCATAATATCTTACCTTAAATGATGATTTAAAACCTCATTTTTTAATTTAAAGG	4385	
Dbb	630 ATCTAGGTGAAGATCNTTTT-GAIAATCNCAIAGCCAAA-TCCCCTAACCGTAGTTTTCG	687	
Cp	4384 ATCTAGGTGAAGATCCCTTTTGATATCTCATGACCAAATCCCTTAACGTGAGTTTTCG	4325	
Dbb	688 TTCACATGAGCGTCAGANCCGGTAGAAAAAT	719	
Cp	4324 TTCACATGAGCGTCAGANCCGGTAGAAAAAT	4293	

4
 RESULT AG004652 721 bp DNA GSS 30-JAN-1998
 LOCUS Homo sapiens genomic DNA, 21q region, clone: 879d12X91, genomic
 DEFINITION survey sequence.
 ACCESSION AG004662
 NID 92826191
 KEYWORDS GSS.
 SOURCE Homo sapiens DNA, clone: 879d12X91.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 721)
 AUTHORS Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.
 TITLE Homo sapiens genomic DNA, chromosome 21q
 JOURNAL Published only in DataBase (1998) In press
 REFERENCE 2 (bases 1 to 721)
 AUTHORS Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.
 TITLE Direct Submission
 JOURNAL Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, Kitasato University, Department of Science, JST
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
 (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
 Fax:0427-78-9561)
 FEATURES
 Location/Qualifiers
 source 1..721
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="879d12X91"
 /map="21q"
 BASE COUNT 195 a 163 c 175 g 179 t 9 others
 ORIGIN
 Query Match 11.98; Score 657; DB 26; Length 721;
 Best Local Similarity 98.7%; Pred. No. 0.00e+00;
 Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4
 Db 32 CATCTTAGCGAIGGCATGCAGATGAAGAAATATGTCAGTCTTGTCTATAATATATATATAT 91
 Cp 4984 CATCTTAGCATGCATGACAGTAAGAAATATATGCTGCTGCTGCTGCTGCTGCTGCTG 4925
 Db 92 AACACTGGCGCCAACTTACTTCTTGACAAAGCATGCGAGGACCGGAAGAGCTAAACCGCTTTT 151
 Cp 4924 AACACTGGCGCCAACTTACTTCTTGACAAAGCATGCGAGGACCGGAAGAGCTAAACCGCTTT 4965
 Db 152 TTGCACAACATGCGGGAATCATGTAACTGCGCTTGTATGTA1103GAAACCGGAGCTGAATGAA 211
 Cp 4864 TTGCACAACATGCGGGAATCATGTAACTGCGCTTGTATGTA1103GAAACCGGAGCTGAATGAA 4805
 Db 212 GCCATACCAACATCGAGCGTGTACATCCACCGCATGCGCTGATGCAATGCAAAACAGCTTGGCG 271
 Cp 4804 GCCATACCAACGACGAGCGTGACATCCGATGCGCTGATGCAATGCAAAACAGCTTGGCG 4745
 Db 272 AAATCTTAACTGGCGAACTACTTACTTCTATCTTATCTTATCTTATCTTATCTTATCTTATCT 331
 Cp 4744 AAATCTTAACTGGCGAACTACTTACTTCTATCTTATCTTATCTTATCTTATCTTATCTTAT 391
 Db 332 GAGCGGATAAAGTTGCAGACACATCTCTGCGTGGGCGGCTTCTGCGTGGCTGGTATT 4685
 Cp 4684 GAGCGGATAAAGTTGCAGACACATCTCTGCGTGGGCGGCTTCTGCGTGGCTGGTATT 4625
 Db 392 GCTGATAAATCTGAGGCTGGTGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 461
 Cp 4624 GCTGATAAATCTGAGGCTGGTGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4565
 Db 452 GATGGTAAGCCCTCCCGTATCGTAGTTTATCTACACGACGCGGAGCTGAGGCACTATGGAT 511
 Cp 4564 GATGGTAAGCCCTCCCGTATCGTAGTTTATCTACACGACGCGGAGCTGAGGCACTATGGAT 4505
 Db 512 CAACGGAATGACAGATCGGTGACATAGGTCCTCACTGATTAAGCCTTGTAAGTACTGCA 570


```
Cp 4654 GCGTGGGCGCTT-CGGGCTGG-CGGGTTATTCCTGATATAAATCGAGATCGAGCGGTGAGCGTG 4597
Db 690 GGTCTC 695
|||||
Cp 4596 GGTCTC 4591

RESULT 10
LOCUS AG003576 696 bp DNA GSS 19-DEC-1997
DEFINITION Homo sapiens genomic DNA, 21q region, clone. p8G4SpN15, genomic
survey sequence.
ACCESSION AG003576
NID g2706702
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: p8G4SpN15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1997) In press
REFERENCE 2 (bases 1 to 696)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato I-15-1, Sagamihara 228, Japan
(E-mail:hattori@rac.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
source
Location/Qualifiers
1..696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="p8G4SpN15"
/map="21q"
BASE COUNT 183 a 165 c 185 g 162 t 1 others
ORIGIN
Query Match 11.3%; Score 624; DB 26; Length 696.
Best Local Similarity 99.1%; Pred No. 0.00e+00.
Matches 680; Conservative 0; Mismatches 0; Indels 5; Gaps 5;

Db 31 GTTTTCTCACCAGAAAGCGTGTGTAAGTAAAGATGCTGAAGATCAGTGGTGCA 90
Cp 5251 GTTTTCTCACCAGAAAGCGTGTGTAAGTAAAGATGCTGAAGATCAGTGGTGCA 5192
Db 91 CGAGTGGTTACATCAACTGGATCTCAACAGCGGTAGATCCTTTGAGAGTTTTCGCCCC 150
Cp 5191 CGAGTGGTTACATCAACTGGATCTCAACAGCGGTAGATCCTTTGAGAGTTTTCGCCCC 5192
Db 151 GAGAGCGTTTCCATGATGAGCATTTTAAAGTTCTGCTATGTGGCGGGTATTATCC 210
Cp 5131 GAGAGCGTTTCCATGATGAGCATTTTAAAGTTCTGCTATGTGGCGGGTATTATCC 5072
Db 211 CGTATTGACGCGGGAAGAGCAACTCGTTCGCCCATACACTATTCTCAGAATGACTTG 270
Cp 5071 CGTATTGACGCGGGAAGAGCAACTCGTTCGCCCATACACTATTCTCAGAATGACTTG 5012
Db 271 GTTCAGTACTCAGAGTCACAGAAAGCATCTTCAGATGGCATGACAGTAGAGAATTA 330
Cp 5011 GTTCAGTACTCAGAGTCACAGAAAGCATCTTCAGATGGCATGACAGTAGAGAATTA 4952
Db 331 TCGAGTGTGTCATAGCATGAGTGAATACAGTCGGCGGCAACTTACTTCTGACACGATC 390
Cp 4951 TCGAGTGTGTCATAGCATGAGTGAATACAGTCGGCGGCAACTTACTTCTGACACGATC 4892
Db 391 GGAGGACCGAAGAGATTAACCGCTTTTGTGCAACATGGGGGATCATGTAACTCGCCTT 450
Cp 4891 GGAGGACCGAAGAGATTAACCGCTTTTGTGCAACATGGGGGATCATGTAACTCGCCTT 4832
Db 451 GATCGTTGGAGTCGGAGTGAGTGAAGAGATACAGTAAAGAGAGCGGTGAGACACCAT 510
```

```
Cp 4831 GATCGTTGGAGTCGGAGTGAGTGAAGAGATACAGTAAAGAGAGAGAGATGAT 4773
Db 511 GGCCTGTAGCAATGGCAACAGCGTTGGCCAAACTATTAACTGGGAACTACTTACTCTAG 570
Cp 4772 G-CCTGTAGCAATGGCAACAGCGTTGGCCAAACTATTAACTGGGAACTACTTACTCTAG 4714
Db 571 CTTCGCGGCGCAACAATTAATAGACTGGATGAGCGGATATAAATGTCAGGACACTCTCIG 630
Cp 4713 CTTCGCGGCGCAACAATTAATAGACTGGATGAGCGGATATAAATGTCAGGACACTCTCIG 4655
Db 631 CACTCGCGGCGCTTTCGCGCTGGGCTGGTIIATTCGTGATATAAATGTCAGGCGGTCAGG-IG 689
Cp 4654 CACTCGCGGCGCTTTCGCGCTGGGCTGGTIIATTCGTGATATAAATGTCAGGCGGTCAGGCTG 4597
Db 690 GGTCTC 695
|||||
Cp 4596 GGTCTC 4591

RESULT 11
LOCUS AG006062 720 bp DNA GSS 14-MAR-1998
DEFINITION Homo sapiens genomic DNA, 21q region, clone. 762015SpN14, genomic
survey sequence.
ACCESSION AG006062
NID g2960469
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: 762015SpN14.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE 2 (bases 1 to 720)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato I-15-1, Sagamihara 228, Japan
(E-mail:hattori@rac.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
source
Location/Qualifiers
1..720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="762015SpN14"
/map="21q"
BASE COUNT 172 a 178 c 170 g 180 t 14 others
ORIGIN
Query Match 11.3%; Score 620; DB 26; Length 720.
Best Local Similarity 97.0%; Pred No. 0.00e+00.
Matches 669; Conservative 0; Mismatches 15; Indels 6; Gaps 6;

Db 35 AGCTCCCTCGTGGCTCTCTCTCTCGGACCTTGGCGGTTATCGGATACCTCTCGCGCTTT 94
Cp 3798 AGCTCCCTCGTGGCTCTCTCTCTCGGACCTTGGCGGTTATCGGATACCTCTCGCGCTTT 3857
Db 95 CTCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAAGTATCTCAGTTCGCTG 154
Cp 3858 CTCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAAGTATCTCAGTTCGCTG 3917
Db 155 TAGCTCGTTTGGTCCCAAGTGGGCTGTGTGCACGAAACCGCGCTTCAGCGGAGAGCTG 214
Cp 3918 TAGCTCGTTTGGTCCCAAGTGGGCTGTGTGCACGAAACCGCGCTTCAGCGGAGAGCTG 4977
Db 215 GCTTATCGGTAAGTATCTCAGTTCGCTGTAAGTATCTCAGTTCGCTGTAAGTATCTCAGT 274
Cp 3978 GCTTATCGGTAAGTATCTCAGTTCGCTGTAAGTATCTCAGTTCGCTGTAAGTATCTCAGT 4037
```



```

/chromosome="21"
/clone="c48B2A23"
/map="21q"

BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 11.2%; Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGATCCGGC 100
Qy 4146 ATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGATCCGGC 4205
Db 101 AAACAACACCCTGCTAGCGGTG-TTTTTTTTGTTCGACACACAGATTACGGCGAGA 159
Qy 4206 AACAAACACACCTGCTAGCGGTGTTTTTTTGTTCGACACACAGATTACGGCGAGA 4265
Db 160 AAAAAGGATCTCAAGAAGATCTTTTGATCTTTTTCAGGGGCTGTGACGCTCAGTGGAA 219
Qy 4266 AAAAAGGATCTCAAGAAGATCTTTTGATCTTTTTCAGGGGCTGTGACGCTCAGTGGAA 4325
Db 220 GAAACCTCAGTTTAAGGATTTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 279
Qy 4326 GAAACCTCAGTTTAAGGATTTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 4385
Db 280 CTTTAAATTAATAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT 339
Qy 4386 CTTTAAATTAATAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT 4445
Db 340 GATAGTTACCAATGCTTAATCAGTGAAGCACTATCTCAGCGATCTGCTATTTCGTTCA 399
Qy 4446 GACAGTTACCAATGCTTAATCAGTGAAGCACTATCTCAGCGATCTGCTATTTCGTTCA 4505
Db 400 TCATAGTTTCGCTGACTCCCGCGTGTAGATCACTACATACAGGAGGCT-ACCATCT 458
Qy 4506 TCATAGTTTCGCTGACTCCCGCGTGTAGATCACTACATACAGGAGGCTTACCATCT 4565
Db 459 GGGCTCAATCTCTCAATGATACGGGAGAGACACACGCTCACGGGCTCCAGATTATCAGCA 518
Qy 4566 GGGCTCAATCTCTCAATGATACGGGAGAGACACACGCTCACGGGCTCCAGATTATCAGCA 4625
Db 519 ATAAACAGGACGCGGAGAGGCGGAGGCGCA-AGTGGTCTGCAACTTTATCGGCTC 577
Qy 4626 ATAAACAGGACGCGGAGAGGCGGAGGCGCAAGGCTGCAAGTGGTCTGCAACTTTATCGGCTC 4684
Db 578 CATCAGCTTATTAATGTTGCGGCGCAAGCTAGAGTAAAGTAGTTTCGCGAGTTAATAGT 637
Qy 4685 CATCAGCTTATTAATGTTGCGGCGCAAGCTAGAGTAAAGTAGTTTCGCGAGTTAATAGT 4742
Db 638 TGGGCAACCTTGTTCATGTTGCTAGAGGATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 697
Qy 4743 TGGGCAACCTTGTTCATGTTGCTAGAGGATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 4801
Db 598 GGTTCATTTTCAGTTCGGGTCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 746
Qy 4802 GGTTCATTTTCAGTTCGGGTCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 4849

RESULT 14
LOCUS AC040352 692 bp DNA GSS 10-JUL-1998
DEFINITION CIT-HSP-2327K21.1F CIT-HSP Homo sapiens genomic clone 2327K21,
genomic survey sequence.
ACCESSION AC040352
NID g3306184
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Pass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Sub,E., Wible,C., Shibuya,H.,

```

Simon,M. and Venter,J.C.
 Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building
 Unpublished (1997)
 Other_GSS: CIT-HSP-2327K21.1F
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..692

/organism="Homo sapiens"
 /note="vector. pBelBAC11; Site_1. HindIII; Site_2:
 HindIII"
 /db_xref="taxon:9606"
 /clone="2327K21"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"

BASE COUNT 158 a 193 c 186 g 155 t

ORIGIN

Query Match 11.0%; Score 604; DB 27; Length 692.

Best Local Similarity 99.8%; Pred. No. 0.00e+00;

Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 87 AGCTTGGGTAATCATGGTCATAGCTGTTTCTGTTGTAATTTGTTATCCCTCACAATT 146

Qy 3273 AGCTTGGGTAATCATGGTCATAGCTGTTTCTGTTGTAATTTGTTATCCCTCACAATT 3332

Db 147 CCACACACATACGAGCGGAAAGCATAAAGTGTAAAGCTTGGGTTGGTTAAATGATGAGT 206

Qy 3333 CCACACATACGAGCGGAAAGCATAAAGTGTAAAGCTTGGGTTGGTTAAATGATGAGT 3332

Db 207 TAATCACAATTAAATGGTTTGGTCTACTTGGGCTTTCGAGTTTGGGTTTGGGTTTGGG 466

Qy 3393 TAATCACAATTAAATGGTTTGGTCTACTTGGGCTTTCGAGTTTGGGTTTGGGTTTGGG 3452

Db 267 CAGCTGCATTAATGATCGGCAAGCGGCGGAGAGAGGCTTTCGAGTTTGGGTTTGGGTTTGGG 526

Qy 3453 CAGCTGCATTAATGATCGGCAAGCGGCGGAGAGAGGCTTTCGAGTTTGGGTTTGGGTTTGGG 3512

Db 327 TCGGCTTCCTGCTCAGTCAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 386

Qy 3513 TCGGCTTCCTGCTCAGTCAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3572

Db 387 GTCACCTCAAGAGCGGTAAATAGGTTATCCAGAAATCAGAGGATATAAGAGAGAGAGAG 446

Qy 3573 GTCACCTCAAGAGCGGTAAATAGGTTATCCAGAAATCAGAGGATATAAGAGAGAGAGAG 3632

Db 447 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGACCGTAAAGAGCGCGGCTTCTGCTGCTGCT 506

Qy 3633 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGACCGTAAAGAGCGCGGCTTCTGCTGCTGCT 3632

Db 507 TTCCATAGGCTCGGCGCGGCTGACGAGCATCACAAAATCAGACGCTCAAGTCAGAGTGG 566

Qy 3693 TTCCATAGGCTCGGCGCGGCTGACGAGCATCACAAAATCAGACGCTCAAGTCAGAGTGG 3752

Db 567 CGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCGCTGGAAGTTCGCTGCTGCTGCT 620

Qy 3753 CGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCGCTGGAAGTTCGCTGCTGCTGCT 3812

Db 627 TCTCCTGTTCCGACCTGCGGCTTACCGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686

Qy 3813 TCTCCTGTTCCGACCTGCGGCTTACCGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3812

US 090207161CP1



Creation date: 28-08-2003
Indexing Officer: FNIGATU - FITSUM NIGATU
Team: OIPEBackFileIndexing
Dossier: 09020716

Legal Date: 18-05-1999

No.	Doccode	Number of pages
1	NPL	8
2	NPL	14
3	NPL	4
4	NPL	10

Total number of pages: 36

Remarks:

Order of re-scan issued on

The first part of the paper discusses the importance of the study and the objectives of the research. It highlights the need for a comprehensive understanding of the subject matter and the role of the researcher in this process. The second part of the paper presents the methodology used in the study, including the data collection methods and the analysis techniques. The third part of the paper discusses the results of the study and the conclusions drawn from the data. The final part of the paper provides a summary of the findings and offers suggestions for future research.

The study was conducted in a systematic and rigorous manner, following the principles of scientific research. The data was collected from a large sample of participants, and the results were analyzed using advanced statistical techniques. The findings of the study are presented in a clear and concise manner, allowing for a thorough understanding of the subject matter.

The results of the study indicate that there is a significant relationship between the variables being studied. This finding is consistent with previous research in the field, and it provides valuable insights into the underlying mechanisms of the phenomenon being investigated. The study also identifies several limitations and areas for future research, which will help to further refine our understanding of the subject matter.

In conclusion, the study has provided a comprehensive overview of the subject matter and has identified several key findings. The results of the study are presented in a clear and concise manner, allowing for a thorough understanding of the subject matter. The study also identifies several limitations and areas for future research, which will help to further refine our understanding of the subject matter.